The Dissertation Committee for Brian Todd Arneson Certifies that this is the approved version of the following dissertation:

ON THE ROLE OF CONCEPT MAPPING ASSESSMENTS IN TODAY'S CONSTRUCTIVIST CLASSROOM

Committee:

J.J. Lagowski, Supervisor

Marvin L. Hackert

Eric V. Anslyn

Dennis A. Kohl

Dee U. Silverthorn

ON THE ROLE OF CONCEPT MAPPING ASSESSMENTS IN TODAY'S CONSTRUCTIVIST CLASSROOM

by

Brian Todd Arneson, B.S., M.A.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

December 2005

Acknowledgements

I would like to extend my thanks and appreciation to Dr. J.J. Lagowski for his dedication in guiding me throughout my graduate experience. Thank you for inspiring me to think big and to press on, regardless of the difficulties in the way.

This work would not have been possible without the programming genius of Andy Yiu, who was instrumental in the development of CMAT.

I would also like to acknowledge the members of the chemical education group – Anna Bergstrom, Michael Elliott, Donna Lyon, Brad Herrick, Deborah Walker, Jessica White, Gloria Brown-Wright, Justin Briggle and Dr. Kent Stewart – for creating an intellectually stimulating experience.

Finally, I would like to dedicate this dissertation to my wife and family. Your love and support along the way has made this experience, while arduous, one that I will look back on fondly. To my parents and sister, thank you for your patience and never ending encouragement. And to my wife, thank you for going through this process with me – helping me when in doubt, encouraging me when in need, and being there always.

ON THE ROLE OF CONCEPT MAPPING ASSESSMENTS IN TODAY'S CONSTRUCTIVIST CLASSROOM

Publication No.______________

Brian Todd Arneson, Ph.D.
The University of Texas at Austin, 2005

Supervisor: J.J. Lagowski

The purpose of this study was to explore the use of concept map assessments in freshman level general chemistry courses. Two strategies were employed in this study. The first strategy involved the creation of a web based concept mapping program capable of scoring concept maps drawn by students. The second strategy involved comparing different methods of scoring concept maps.

Students enrolled in web based general chemistry course drew concept maps using the web based Concept Map Assessment Tool, CMAT. The reliability of the automated scoring in the CMAT program was tested by scoring the concept maps created in the CMAT program by hand. The results of the study indicated that scoring concept maps by hand was the same as the automated scoring of concept maps in the CMAT program.

Two characteristics of concept maps serve as the basis for scoring methods. The relational character of a concept map is defined as the correctness of the propositions in the concept map. The structural character of a concept map is defined as the key features of the map, such as branches, long chains or intersecting points. The scoring method used in the CMAT program scores the relational aspects of a concept map. In this study, a second relational scoring method was used to score the concept maps drawn by students using the CMAT program, and the two sets of scores were compared. A novel structural scoring method, the Structural Complexity Index (SCI), was developed compared to the relational scoring approach of the CMAT program. The results of this study found the two relational scoring methods to score concept maps similarly under certain conditions. The SCI was found to produce a different score for concept maps than the relational scoring method employed by CMAT.

Table of Contents

List of Figures

List of Tables

Chapter 1. Introduction

"I contend that the greatest impediment to progress regarding the chemistry curriculum is our lack of ability to measure whether we have accomplished increased student learning of the kind we really want."
–J.W. Moore [1]

INTRODUCTION

The quantification of learning is currently en vogue, with interest in the assessment of learning growing even beyond academia. Emphasis on research into the kinds of assessment used to quantify learning has increased both the stakes invested in testing and the debate on the quality – or worthiness – of assessment methods. The use of high stakes testing in K – 12 schooling for a multitude of purposes has made the inadequacies of traditional multiple choice tests more obvious [2 – 4]. A curriculum geared towards sufficient proficiency on a test, instead of mastering the conceptual underpinnings, will produce a student population ill-suited for learning in university environments as well as for the necessary life-long learning experiences that lead to success [2, 5], unless of course the university environment adapts to the incoming students. This dissertation addresses the issue of assessment in undergraduate chemistry education. The goal of the dissertation is to contribute to the discussion on assessment with the expectation of bringing about a more effective means of measuring the quality of student learning.

STATEMENT OF PROBLEM

Moore's statement [1] made more than a decade and a half ago was a review of the current state of chemical education at the end of the 1980's. At the time, Scantron technology dominated the university classroom, as it still does, and Moore's challenge was to change the assessment status quo. Another harbinger for change comes from educational psychologists; from the end of 1980's through the present, constructivism has emerged as a guiding educational theory [6]. Constructivism calls for an assessment centered instruction. However, the nature of assessment within the constructivist paradigm differs significantly from the traditional testing environment [7 – 9]. It is under the constructivist paradigm that this dissertation addresses the assessment of learning via an alternate assessment method – namely web based concept mapping.

The most common assessments in today's classrooms are much the same as they were 50 years ago, multiple choice tests. While multiple choice tests have some positive aspects, they have been widely criticized (for many years!) as a method to measure of student learning [2]. Assessment, in terms of accepted educational language, is the collection of information that describes a student's level of understanding. The ability to measure student learning is dependent upon the information collected from students. The shear number of students in entry level science courses has lead to reliance on easy to score methods of assessment. Especially using a paper and pencil medium; multiple choice questions and automated scoring is viewed as the only practical method to deal with the numbers of students involved.

> "Many examinations, and not just in chemistry courses, appear to ignore the technology currently available to (and used by) students. The irrelevance of such examinations becomes obvious when one realizes that a good pocket calculator could pass them if only it know which keys to press." -J.W. Moore [1]

The realization that the medium has driven the assessment – and hence, the educational – choices in the classroom has been met with a flood of web based assessment tools in recent years [10]. These on-line assessment tools have taken the standard assessment off of the paper and placed it on the computer screen. The types of assessments have increased to include more than just multiple choice questions. With the increased flexibility in creating on-line assessment tools, the capability exists to collect more information used to assess a student's learning; therefore, the onus is on the researcher to determine what is the "learning we really want", as Moore put it, to assess. Bloom's Taxonomy of cognitive factors (Figure 1) provides a guide to define learning outcomes in terms of cognitive abilities required of a student [11].

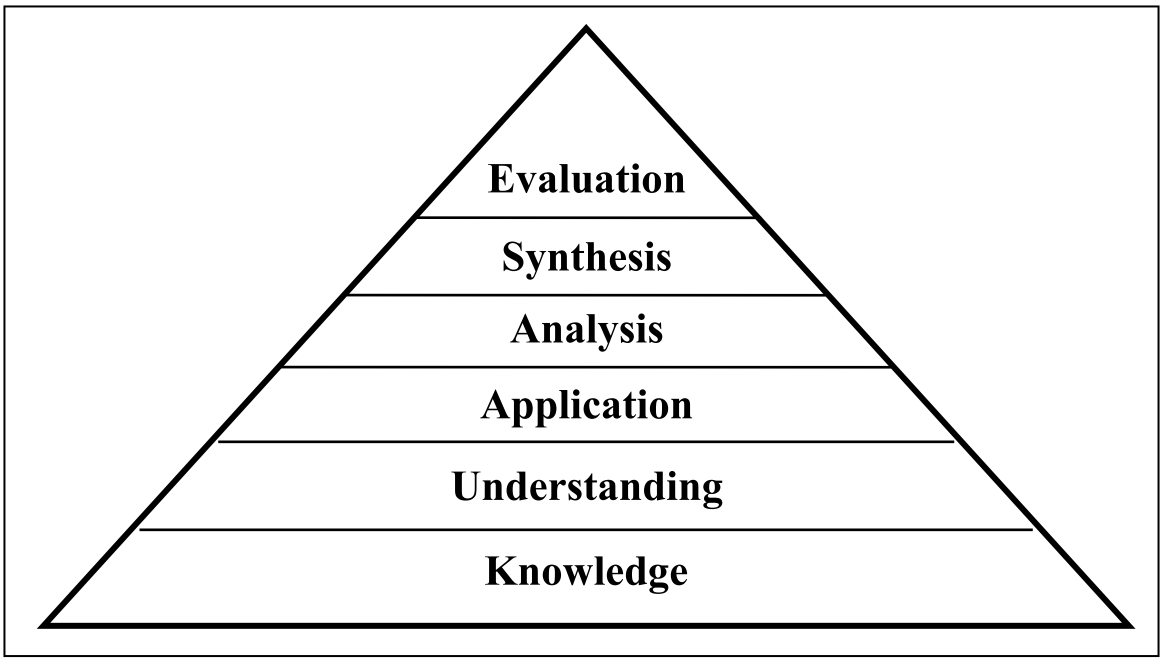


Figure 1 Bloom's Taxonomy of the Cognitive Domain.

The seven levels in Bloom's Taxonomy are listed in terms of increasing complexity, starting with knowledge at the lowest end and evaluation at the highest. Within the context of Bloom's Taxonomy, each of the levels can be described by what is asked of the student. The "Knowledge" level requires students to recall information such as a definition. The "Understanding" level requires students to comprehend the meaning, or be able to explain in their own terms, of a concept or set of instructions. The

“Application” levels requires students to apply what is known to a new situation, whether it is in or out of the classroom. The “Analysis” level requires students to separate complex materials or concepts into component parts or to interpret graphical representations; analysis also includes the ability to decide between fact and inference. The “Synthesis” levels requires students to construct a pattern or structure from simpler concepts or material; the organization of a diverse array of ideas into meaningful sets of information. The “Evaluation” level requires students to judge the value of an idea or product, whether it is their own work or from someone else.

Bloom’s Taxonomy has become the standard upon which much of the current literature uses to determine the quality of an assessment. An assessment that includes all of the levels of Bloom’s Taxonomy is judged to be *“better”* than an assessment with just a few of the levels addressed. The current multiple choice instruments address only the lower order skills from Bloom’s Taxonomy. Higher order cognitive skills, often abbreviated HOCS, refers to the higher levels of Bloom’s Taxonomy and are often missing in traditional assessments. The influence of the constructivist paradigm on the National Science Education Standards [6] is apparent in the call for an assessment centered curriculum that touches on all levels of cognitive ability. The assessments within this environment take on both a formative and summative role. Formative assessments are used to identify student misconceptions and are used to guide future learning. Summative assessments are used as major benchmarks of achievement within the course structure. To incorporate the different levels of cognitive abilities and different roles of assessment, a variety of assessment types is recommended. Indeed, the reliance upon a sole type of assessment has been shown to not provide an adequate overall picture of student learning [5].

PURPOSE OF THE STUDY

A goal of this dissertation to contribute to the body of knowledge concerning assessment in today's classrooms within the constructivist paradigm. Today's classroom is a demanding environment where student to teacher ratio is extremely high, resulting in the need for electronic tools to assist teachers in the assessment of learning by their students. One assessment tool that has been reported for years as being an effective formative assessment technique is concept mapping [12]. However, the nature of the concept mapping process precludes it from being readily integrated into classrooms [13]. The time intensive nature of scoring concept maps has created a difficulty in integrating concept map assessments into the curriculum. In addition to the length of time necessary for scoring, there are many different scoring rubrics reported. There are, also, some concerns about the validity of scoring concept maps [14 – 17]. This study investigates the use of a web based concept mapping system that assists students in creating concept maps equally and that automatically scores concept maps. In addition to the scoring algorithm used in the web based concept mapping system, two other scoring techniques are investigated. We expect that the results of this study will provide useful information concerning the worthiness of web based concept mapping and the information that concept map scores can provide in the assessment of learning.

GUIDING QUESTIONS

To address concerns raised in using concept maps, three questions are addressed:

- *Can a web based concept mapping program score concepts accurately?* This dissertation focuses on creating a web based environment where students can create concept maps incorporating an automated scoring algorithm for immediate feedback. The validation of the electronic scoring algorithm, compared to scoring the same maps by hand, will provide necessary information on the ability to use web-based concept mapping as an assessment tool. In addition to the creation of the web based concept mapping environment, the scoring of concept maps is investigated to determine how that information can be used for the assessment of learning.

- *Do different scoring schemes produce different scores for the same concept map? If so, is there a reason for the difference?* Concept maps can be imagined to exhibit a structural nature and a relational nature. The structural nature of a concept map consists of how different concepts are arranged into a hierarchical structure. The relational nature consists of the right or wrong connections between individual concepts. This study will investigate the use of two different relational scoring techniques and compare the results. In addition to the relational scoring methods, a structural scoring technique is developed and used to score the concept maps. This method differs significantly from the scoring method of electronic scoring used in the web based concept mapping tool developed.
- *How does the concept map scoring techniques fit into the constructivist paradigm?* A review of educational theories provides a framework for using concept maps as measures of *meaningful learning* as defined by Ausubel [18]. The implications of using web based concept mapping as it pertains to the assessment centered nature of learning in the constructivist paradigm is discussed.

Chapter 2. Literature Review

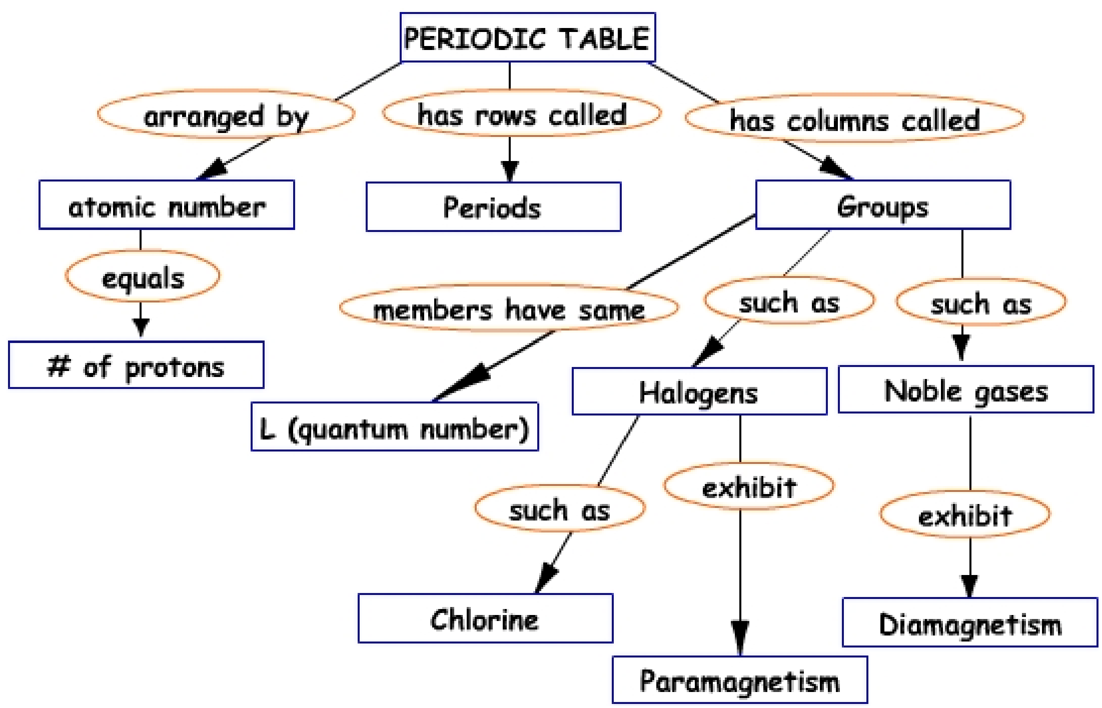
The idea of constructivism as an educational theory is mostly attributed to Piaget [19] and Ausubel [18], however, earlier educational psychologists, such as Vygotsky [20], proposed models of learning consonant with the learning theories of the latter researchers. The basic tenet of constructivism is the learner constructs their own understanding of knowledge based upon previously existing knowledge they possess.

Ausubel's fundamental assumption of cognitive learning is encapsulated in the statement:

> If I had to reduce all of educational psychology to just one principle, I would say this: The most important single factor influencing learning is what the learners already know. Ascertain this and teach accordingly. [18]

More recently, many authors have emphasized the use of constructivist educational theory in chemical education [9, 21 – 25]. The implications of teaching within the constructivist paradigm means that students do not acquire the exact knowledge of the instructor, rather students learn according to how their own idiosyncratic experiences have prepared them to learn. This rather maddening outcome leads to the identification of misconceptions, or alternate conceptions, as an important factor in the learning process [25 – 27]. The role of assessment in the constructivist paradigm is to diagnose misconceptions for both student and teacher [9, 25]. The importance of eliciting existing knowledge – and how it is structured – has become an important aspect of the assessment of school learning [28]. Joseph Novak and Bob Gowin coined the term 'concept map' as the tool used to measure conceptual change in their studies with elementary school science students. Concept maps were used to represent a graphical manifestation of the knowledge structure a student possessed and how the knowledge structure was altered after instruction [26].

The basic unit of a concept map is the proposition. A proposition is a concept – linking phrase – concept statement (see Figure 2). Propositions are arranged in a hierarchy such that the most general, or abstract, concept is placed highest in the concept map, and more specific concepts are below it. Propositions that consist of examples are the lowest in the hierarchy. Propositions that link two separate chains within the hierarchy are called cross-links, and Novak [26] uses as an example of a reconciliatory process that is evidence of meaningful learning.



Concepts are contained within boxes and linking phrases are contained within ellipses. Each arrow represents a proposition, a concept–linking phrase–concept statement.
Note: The term Halogens represents the *elements* in Group VIIA.

Figure 2 Example of a concept map

Other methods used to represent knowledge structure, such as semantic networking and fill-in-the-structure [28], have similar aspects to concept mapping. Concept mapping has been more prominently used in science education [27] and has gained a wide audience. A review of concept mapping literature in 1993 shows concepts maps have been used in the classroom for a variety of activities [12]. The use of concept mapping as an assessment has been shown with mixed results [17] and there are issues that must be addressed in using concept maps as an assessment tool [14 – 16].

Despite the acceptance that concept mapping is an important and beneficial educational tool, the widespread use of concept mapping as an assessment tool has been relatively low [13]. The main hindrance in utilizing concept maps is the difficulties for instructors to evaluate the concept maps in a timely and reliable manner [13, 17]. A solution to this problem is the computerization of the concept mapping activity. For several years, the construction of concept *via* a computer interface has been available from such software as Inspiration®. Many studies have confirmed that students find constructing concept maps via computer interface easier and more beneficial than the traditional approach of pencil and paper [29, 30]. However, web based concept mapping packages capable of evaluating a concept map once it is create are in their infancy [31–35].

The ability to describe a learner's cognitive structure through concept mapping has led to the use of concept maps as assessment tools. The use of concept maps as an assessment then required the development of a scoring rubric to judge the quality of the constructed map. Novak and Gowin [26] originally proposed that concept maps should be scored based upon the number of identifiable features included in the concept map. These features include: correct propositions, appropriate hierarchical arrangement, amount of integration between disparate concept branches (cross-links), and listing specific example. The weighting of the scores for each of these features reflected Novak and Gowin's own idiosyncratic view of what a concept map represents (Figure 3). For example, if a correct proposition is scored as value of 1, a cross-link is scored as a value of 10 because in Novak and Gowin's view, the cross-link is an expression of *integrative reconciliation*, an important sign of meaningful learning. Also, examples are explicitly separated from other propositions. Examples are concrete references which can be numerous and therefore outweigh the fewer, more meaningful propositions in the concept map. Therefore, a possible variant of the scoring scheme could score examples as one-half the value of a proposition. Novak and Gowin encouraged researchers to experiment with variations on their scoring scheme as they were not convinced of its value [26]. When scoring concept maps using Novak and Gowin's approach, a sum total of the propositions, hierarchies, cross-links and examples is calculated and used to represent the quality of a student's concept map.

$$\begin{array}{lcl} \text{Propositions (if valid)} & = & 1 \times n \\ \text{Hierarchies (if valid)} & = & 5 \times n \\ \text{Cross links} & = & 10 \times n \\ \text{Examples} & = & 1 \times n \\ \hline \text{Total} & = & \end{array}$$

Figure 3 Concept map scoring algorithm from Novak and Gowin.

Many researchers have used variations on Novak and Gowin's original scoring scheme [14, 36]. Scoring schemes are dependent upon how the concept map is created. One class of scoring is used primarily in concept mapping activities where students generate a concept map *de novo* with no assistance. Stuart [40] proposed using a scoring scheme that is similar to Novak and Gowin's in that hierarchy, correct propositions and cross-links are scored. However, Stuart also included a branching element, number of correctly used technical terms, and a category called 'general – to – specific.' Another caveat in Stuart's scoring of concept maps that set this method apart from Novak and Gowin's is keeping the component scores separate. A second class of scoring schemes reduces the emphasis Novak and Gowin placed on assessing knowledge structure and instead increases the emphasis on propositional correctness. Scoring schemes in this class vary between including a directionality requirement to the proposition [17] or no directionality component [37].

$$\frac{\#\text{ correct (linking phrases)} - \#\text{ wrong or non-informative}}{\text{total }\#\text{ of connections}} \times \text{points possible}$$

Francisco and Nakleh's scheme for scoring concept map is dependent entirely on the correctness of the propositions formed. The purpose of using this scoring system was to evaluate a student's understanding of a chemistry laboratory activity.

Figure 4 Francisco and Nakleh's scoring algorithm for concept maps.

The advantage gained with these scoring schemes has been in reliability of scoring with different evaluators [14, 17, 37]. A higher correlation between concept mapping scores and traditional examinations are seen when using scoring schemes that emphasize propositional correctness [14, 17]. A third class of concept map scoring exists and stands apart from the others in that it is not based upon Novak and Gowin's original scheme at all. These scoring methods look at the structure of the map [38, 42] irrespective of the propositional correctness. By classifying the structural elements of the concept map into categories such as chain-like, spoke-like or network-like [38] these scoring schemes aim to relate the complexity of the student's cognitive structure (see Figure 5).

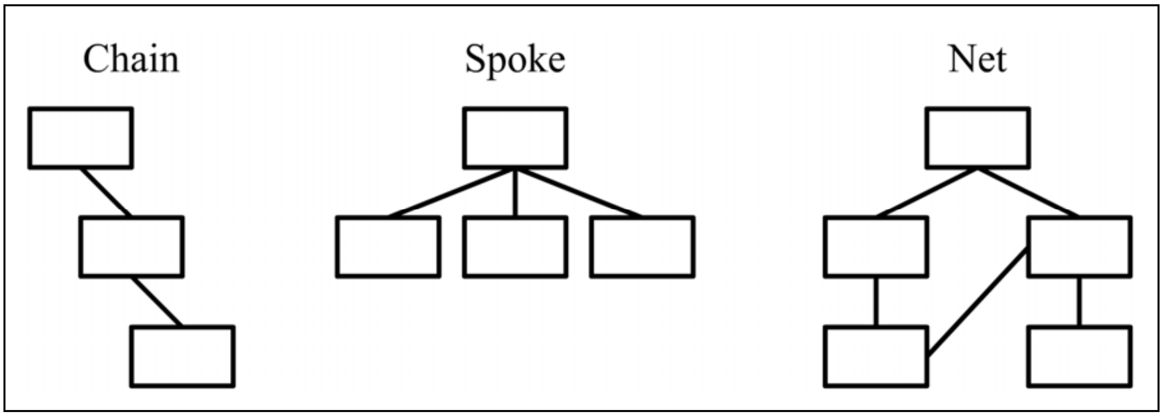


Figure 5 Characterization of concept maps based upon structural features termed by Kinchin as chain, spoke and net.

Traditionally, scoring concept maps has been a difficult and time-consuming process. This fact along with many educators' lack of experience with concept mapping has hindered their wide-spread use as an assessment tool in most classrooms [13]. With all of the variation present in concept map scoring, Ruiz-Primo and Shavelson called for concept map assessments to have a theoretical background [14].

Chapter 3. Methodology

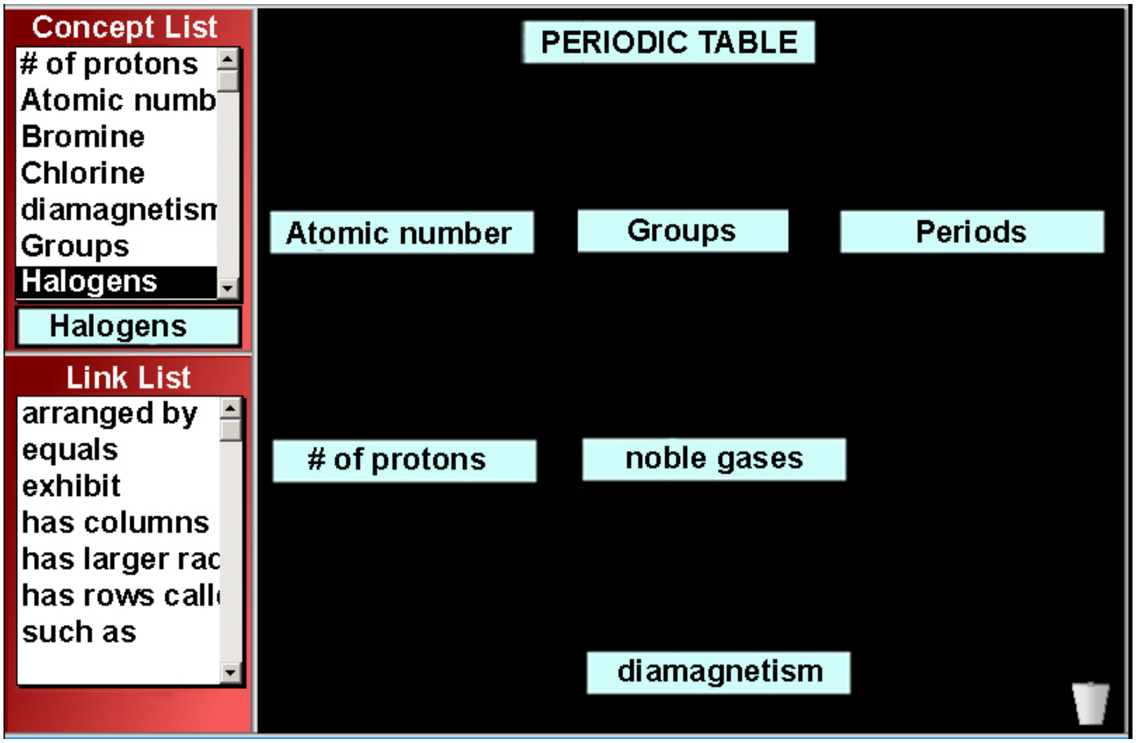
In this chapter, the methodology of the study is described. The study encompasses 3 different experiments which test a total of 4 separate hypotheses. Also included in this chapter are descriptions of the components of the study. These components include the web based concept mapping program – and associated concept map scoring scheme, a concept map scoring scheme for structural complexity, the web based freshman general chemistry course where the study was conducted, and the three concept mapping assignments.

CONCEPT MAP ASSESSMENT TOOL (CMAT)

Concept mapping has been used as an assessment tool by placing the construction and scoring of concept maps into a web-based environment. However, the automated scoring of concept maps has been slow in developing. The first attempts to score a concept map automatically used a fill in the blank concept map where a completed map was presented to the learner with some concept labels missing [30]. This approach was essentially "fill in the blank" and it could not lead to any pedagogical advantage over what we now do. More recently, much more advanced automated scoring has been incorporated into these web-based systems [31 – 35]. Our own development of a web-based concept map authoring and scoring environment led to the creation of CMAT, Concept Map Assessment Tool, that we have used in a web-based general chemistry course [39] during the course of the study presented here.

Basic operation overview

The Concept Map Assessment Tool, CMAT, consists of two components, an interface component accessible with a web browser and a CGI script that controls scoring and record keeping functions. The CMAT interface program was created using the Macromedia Director® authoring software. The Macromedia Shockwave Plug-in is required for to run the CMAT program within a web browser. The CMAT interface allows students to construct a map from a list of concepts and linking words and submit their map for scoring. CMAT was designed to mimic the note card concept mapping technique described by Novak and Gowin [26]. For each concept mapping assignment, the instructor creates a list of concepts and a list of linking phrases. When a student selects a concept mapping assignment, CMAT retrieves the appropriate concept list and linking phrase text files and displays them in the CMAT authoring environment (Figure 6). The concepts and linking phrases are put into separate, scrolling lists that appear in the CMAT authoring environment.

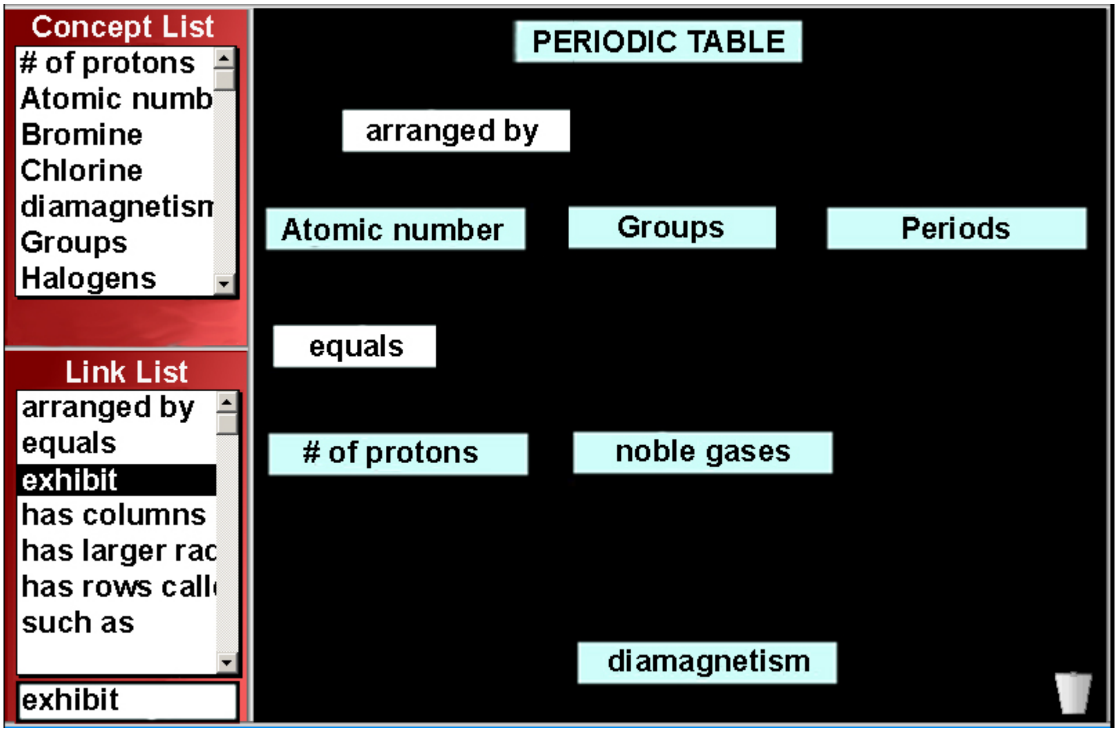


Concepts available for a particular concept mapping assignment are listed in the scrolling text box titled "Concept List" and linking phrases are listed in the scrolling text box titled "Link List." Clicking on a concept in the Concept List creates a light blue box containing the concept. The light blue box can be dragged to any area of the drawing area (black window). Dragging a concept into the trash can on the lower right of the screen removes the concept from the drawing area.

Figure 6 The CMAT authoring environment during concept placement.

Clicking on a concept will create a colored box containing the concept; each concept can be used only once. This box can be dragged around the CMAT drawing area. The process of selecting a concept and dragging into place is repeated until a student does not want to use anymore concepts; only one occurrence – or instance in terms of graphical programming language – of a concept is permitted in the CMAT drawing area.

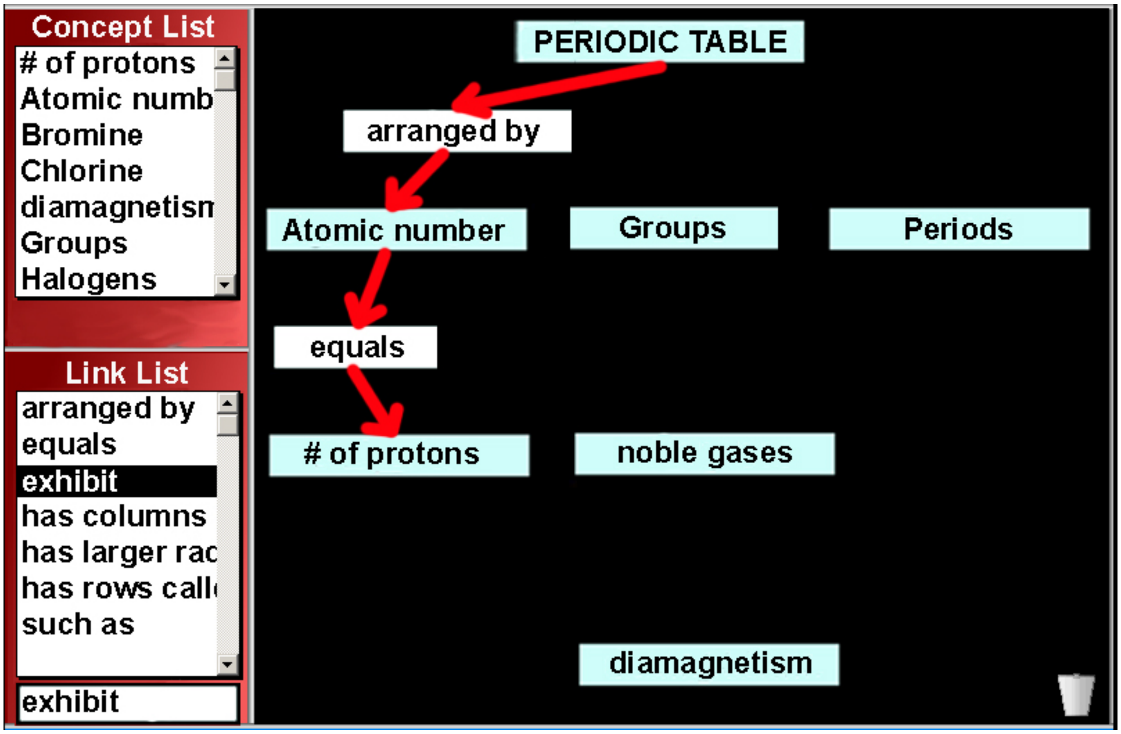
A similar procedure is then used to place linking phrases between concepts that a student wants to connect to form a proposition (Figure 7).



Clicking on a linking phrase in the Link List creates a white box containing the linking phrase. The white box can be dragged to any area of the drawing area (black window). Dragging a linking phrase into the trash can on the lower right of the screen removes the linking phrase from the drawing area. More than one instance of a linking phrase can be placed into the drawing area.

Figure 7 The CMAT authoring environment during linking phrase placement.

More than one occurrence of a linking phrase can be placed in the CMAT drawing area. After the concepts and linking phrases are arranged, they can be connected together in a concept – linking phrase – concept sequence (Figure 8).



Propositions are created by clicking first on a concept then a linking phrase and then a second concept. An unlimited number of propositions can be created.

Figure 8 The CMAT authoring environment during proposition creation.

A student can remove the linking arrows by clicking on appropriately labeled buttons. When the student is completed with their concept map, clicking the submit button sends the concept map information to a CGI script on the web server. The CGI script scores the submitted concept map, stores the student response in a text file (Figure 9), and then returns the score to the CMAT interface for the student to view.

Student 19	90
PERIODIC TABLE,arranged by,Atomic number
PERIODIC TABLE,has columns called,Groups
PERIODIC TABLE,has rows called,Periods
Atomic number,equals,# of protons
Halogens,exhibit,paramagnetism
noble gases,exhibit,diamagnetism
noble gases,members have same,L (quantum number)
Halogens,such as,Chlorine
Groups,such as,Halogens
Groups,such as,noble gases

The student is identified along with the score generated by CMAT on the top line of the entry. Each proposition in the concept map is listed on a separate line. The concept in all capital letters is given to the students as the top-most concept in the hierarchy.

Figure 9 Example text file entry of a concept map drawn in CMAT.

CMAT scoring algorithm

The concepts and linking phrases included in the concept mapping activities were carefully chosen to (logically) limit the number of possible propositions. The scoring of concept maps in CMAT was based upon the correctness of each proposition. The CGI script analyzed each proposition formed in the concept map by comparing it against a master list of acceptable propositions created by the instructor. Positive points are awarded to a proposition contained in the master list and points were subtracted for propositions not included in the master list. To create the master list of propositions, the instructor creates a concept map, used as a master map, through the CMAT administrator interface. The instructor assigns a point value to each proposition created in the master map. In this manner, an instructor can assign a value of zero to propositions that were not necessarily incorrect, but did not add value to the concept map. The final score assigned to a concept map is calculated by first taking the total points accrued for the proposition divided the total possible points from the master map then multiplying by 50 and adding 50 points (Equation 1).

$$\text{CMAT score} = \frac{\text{Student's proposition points}}{\text{Master map proposition points}} \times 50 + 50 \qquad \text{(Equation 1)}$$

The concept map score is essentially on a 50 to 100 point scale, unless a student made more incorrect propositions than correct propositions.

STRUCTURAL COMPLEXITY INDEX

Quantitative measures of ‘correctness’ within a concept map [17, 26, 37] are more prominent than qualitative analyses [38, 40]. However, trying to quantify the structural aspects or knowledge structure exemplified in a concept map drawn by a student is just as important as the relational correctness of the propositions [26]. It is the changes in knowledge structure that Ausubel uses to describe learning processes [18, 41]. Therefore, a structural complexity index has been created to provide an indication of the complexity of the knowledge structure displayed in a concept map.

A completed concept map looks like a network of concepts and linking phrases. There are many other disciplines other than education that employ a complexity analysis of a network. Computer programming, for instance, maps the command structure of programs. These command structures are then analysis for circular complexity to determine whether a program is prone to failure. In this application of complexity analysis, the number paths through a program are calculated. As more and more decision or branching points are added to the program, the more complex the command structure. If circular complexity is too high for a particular application, the program is judged to be prone to failure and should be simplified. A command structure that has a low circular complexity is not robust enough to handle differing conditions and probably is not suitable. The values at which decisions about complexity are made within a program, however, differ from application to application. Once standards are set within a particular application, the circular complexity value helps evaluate the quality of a program.

Considering concept maps are a reflection of knowledge structure, we must define what the sub-structures within a concept mean as it pertains to chemistry education. Kinchin has categorized three sub-structures of concept maps commonly found in student drawn concept maps (Figure 5, Chapter 2). The chain structure represents sequential learning, whether or not the sequence is coherent. Chains can also represent procedural processes relevant to a particular concept. A spoke structure represents the stimulus-response type learning often associated with multiple choice learning. A student wants to relate everything to a main concept with little regard for integration or procedural processes. A net structure represents an integrated knowledge structure that typifies meaningful learning. Relationships are drawn between concepts that differ from the relationships between the overarching topics.

Assumptions behind the structural complexity index

The structural complexity index created in this study reflects the view that the net type structural shows the most evidence of meaningful learning. Therefore, changes in a concept map that increases the amount of net structure should, in turn, increase the structural complexity index. Within chemistry education, there are examples of procedural processes that are necessary to understand if a student is to be successful; therefore, increases in chain length should positively affect the structural complexity index. A concept map with many branches but little hierarchical levels, or interconnections, does not show a deep understanding. With these assumptions, or values, the structural complexity index was created.

The structural complexity index (SCI) is calculated using the following four properties:

1. Propositions (P) – the number of propositions contained in the concept map
2. Branches (B) – the number of concepts that have more than one proposition originating from itself
3. Chains (C) – the number of independent paths through the concept map
4. Average chain length (Cavg) – the average number of propositions contained in the chains

$$SCI = (C_{avg} \times P) + (B \times C) \qquad \text{(Equation 2)}$$

Example of Structural Complex Index scoring

Four example concept maps, each with 10 concepts, are shown in Figure 10 to illustrate the SCI scoring approach. To make the example easier to understand, linking phrases are omitted from the concept maps. In Tables 1 – 4, the SCI is calculated for each example concept map found in Figure 10.

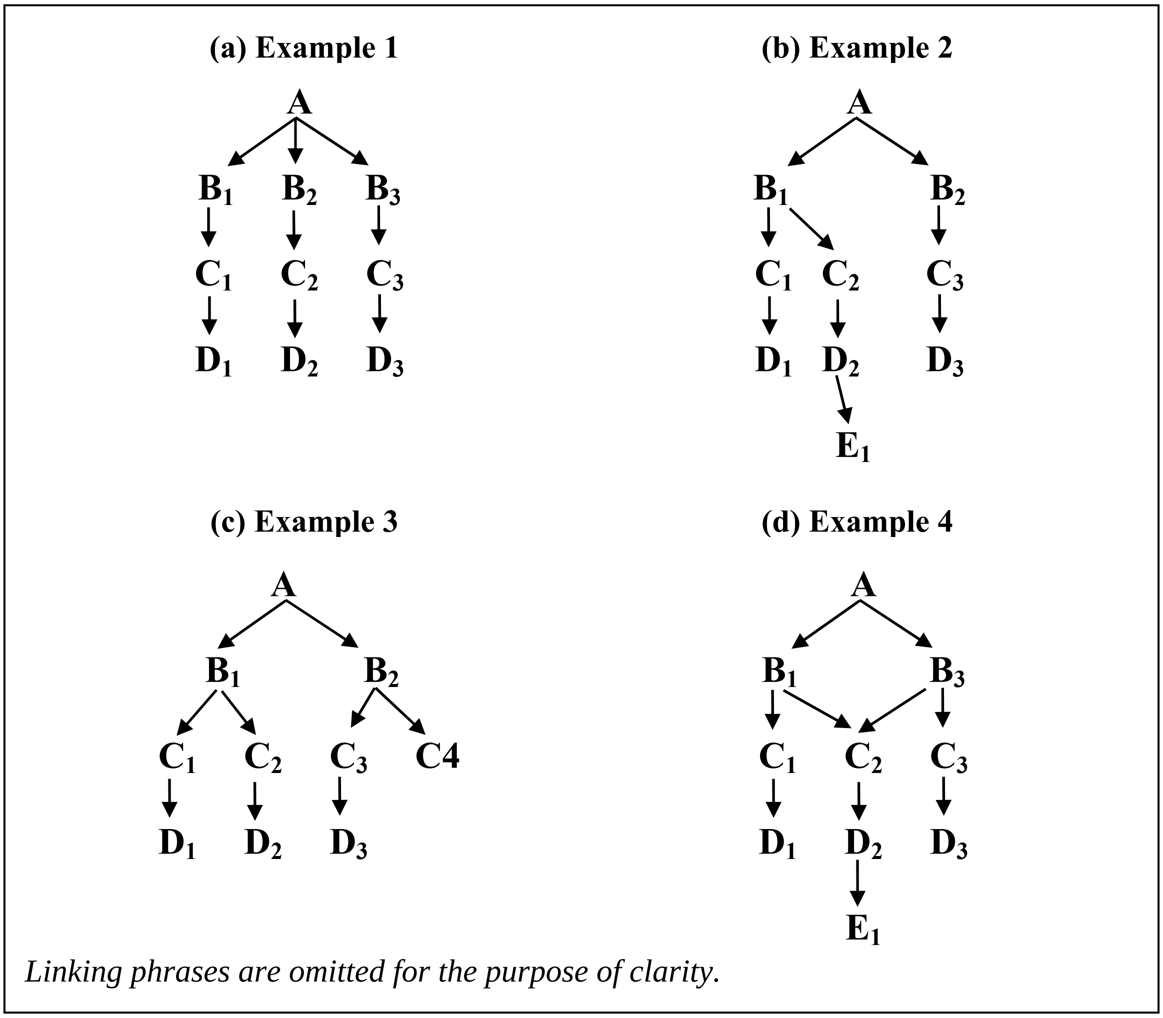


Figure 10 Four concept maps with different structural features used to demonstrate the SCI scoring approach

Table 1 Structural Complex Index for example concept map in Figure 10 (a).		
Property	**Calculation**	**Value**
Propositions		9
Branches	A	1
Chains	$A \rightarrow B_1 \rightarrow C_1 \rightarrow D_1$ $A \rightarrow B_2 \rightarrow C_2 \rightarrow D_2$ $A \rightarrow B_3 \rightarrow C_3 \rightarrow D_3$	3
Average chain length	(3 + 3 + 3) ÷ 3 =	3
Structural Complexity Index	**(3 × 9) + (1 × 3) =**	**30**

Table 2 Structural Complex Index for example concept map in Figure 10 (b).		
Property	**Calculation**	**Value**
Propositions		9
Branches	A, B_1	2
Chains	$A \rightarrow B_1 \rightarrow C_1 \rightarrow D_1$ $A \rightarrow B_1 \rightarrow C_2 \rightarrow D_2 \rightarrow E_1$ $A \rightarrow B_2 \rightarrow C_3 \rightarrow D_3$	3
Average chain length	(3 + 4 + 3) ÷ 3 =	3.33
Structural Complexity Index	**(3.33 × 9) + (2 × 3) =**	**36**

Table 3 Structural Complex Index for example concept map in Figure 10 (c).		
Property	**Calculation**	**Value**
Propositions		9
Branches	A, B_1, B_2	3
Chains	$A \rightarrow B_1 \rightarrow C_1 \rightarrow D_1$ $A \rightarrow B_1 \rightarrow C_2 \rightarrow D_2$ $A \rightarrow B_2 \rightarrow C_3 \rightarrow D_3$ $A \rightarrow B_2 \rightarrow C_4$	4
Average chain length	(3 + 3 + 3 + 2) ÷ 4 =	2.75
Structural Complexity Index	**(2.75 × 9) + (3 × 4) =**	**36.75**

Table 4 Structural Complex Index for example concept map in Figure 10 (d).

Property	**Calculation**	**Value**
Propositions		10
Branches	A, B_1, B_2	3
Chains	$A \rightarrow B_1 \rightarrow C_1 \rightarrow D_1 \rightarrow E_1$ $A \rightarrow B_1 \rightarrow C_2 \rightarrow D_2$ $A \rightarrow B_2 \rightarrow C_2 \rightarrow D_2$ $A \rightarrow B_2 \rightarrow C_3 \rightarrow D_3$	4
Average chain length	$(4 + 3 + 3 + 3) \div 4 =$	3.25
Structural Complexity Index	**$(3.25 \times 10) + (3 \times 4) =$**	**44.5**

Figure 10 (a) is the most basic example of the four concept maps and the SCI of 30 is the lowest example. With subsequent branching points are introduced as in Figures 10 (b) and (c), the SCI increases. However, with only 10 concepts in each map, the addition of a branching point in Figure 10 (c) reduces the average chain length. Therefore, the example in Figure 10 (c) is scores only marginally higher than the example in Figure 10 (b). The example in Figure 10 (d) contains a networked structure at the C_2 concept. Therefore, this example has an additional proposition the other examples do not, which greatly increases the SCI. Also, even though only three terminal concepts exist in Figure 10 (d), four chains are present because of the networked structure; the chain $A \rightarrow B_1 \rightarrow C_2 \rightarrow D_2$ is different than the chain $A \rightarrow B_2 \rightarrow C_2 \rightarrow D_2$. These four examples highlight the principles of calculating the SCI for concept maps. The difference in SCI between the different examples shows the increases in complexity and allows the categorization of the concept maps.

Summary of Structural Complexity Index

The SCI created during this study can be used to differentiate between subsequent concept maps that display difference in structural characteristics. The technique of assigning a value of structural character to a network is done in other fields [43]. As in other applications, structural scoring by itself does not necessarily provide the ultimate measure of goodness displayed by the networked structure, but it is helpful in making assessment judgments.

CH 301WB – A WEB BASED FRESHMAN GENERAL CHEMISTRY COURSE

The first semester general chemistry course at The University of Texas at Austin is CH 301. The content of CH 301 is the material found in the first twelve chapters of "General Chemistry" by Whitten, Davis, Peck, and Stanley. CH 301wb is a web based equivalent of CH 301. CH 301wb is a pseudo-self-paced course with the pace being dictated by the placement of the chapter assignment deadlines. This general structure breaks up the subject to be covered into smaller pieces upon which the students can concentrate; this strategy ensures the students' efforts to learn with the component features is spread more-or-less evenly over the course of a standard semester and is not allowed to accumulate at the end of the course, a condition that often is found in conventional self-paced courses and that leads to disastrous consequences for immature learners. The level of achievement expected for students taking CH 301wb is the same as that for the standard course.

CH 301wb incorporates many of the familiar features of standard face-to-face courses and it also incorporates some additional features. The components of CH 301wb are the following:

- Lectures (in a multimedia format)
- Crossword Puzzles
- Homework
- Quizzes
- Examinations

- Help sessions, on-line, and face-to-face

These elements of the course are presented in a seamless environment, each contributing its particular pedagogic characteristics to assist students with learning the CH 301-oriented materials.

CH 301wb component descriptions

Lectures

The subject matter in the textbook, chapters 1-12, is divided into 54 multi-media lectures that are distributed across the chapters in the required textbook. The lectures incorporate an oral component, demonstrations of appropriate chemical phenomena, and animations of microscopic interpretation of phenomena and concepts where appropriate. The multimedia presentations can be stopped by the student at any point and replayed.

In general, the strategy of the lecture presentations involves the use of evidence to provide the support for the general principles—the approach is inductive. The lectures also use a “spiral” approach to the subject, that is, a given phenomenon (or concept) is revisited several times during the course to provide a greater depth of insight into, or an expansion of the subject.

Crossword Puzzles

The material in each chapter has an associated, required crossword puzzle that is graded. The crossword puzzle is accessed as a Java Applet® and is completed, submitted and graded on-line.

Homework

The material in each chapter has an associated, required homework set that is graded. For the most part, the questions are taken from the chemistry database associated with the Physics Homework Service.

Quizzes

Short quizzes are interspersed throughout the lectures at appropriate places in the material presented. The quiz questions are taken from the chemistry database associated with the Physics Homework System; when necessary or appropriate, other questions are used that have been designed for CH 301wb.

Examinations

Hour examinations, incorporating questions from the chemistry data base associated with the Physics Homework System and the American Chemical Society General Chemistry test bank, are given at chronologically appropriate points in CH 301wb. The placement of the hour examinations in the flow of the course is a key factor in keeping the students "on pace" in this course. The questions for the final examination are also taken from the chemistry databases.

On-line Discussions

On-line, synchronous office hours accommodate most of the content-oriented questions students might have. An asynchronous bulletin board also serves as a communication tool for students and TA's.

Scheduled Class Meetings

A class meeting time in a classroom (a secure computer laboratory) is scheduled to accommodate the number of students who are enrolled. This provides a single time slot in which students can meet, face-to-face with an instructor at least once a week should that be necessary. This time slot is used as a controlled environment in which examinations can be given. The initial orientation to the course is done during this scheduled class meeting.

CONCEPT MAPPING ASSIGNMENTS

Three concept mapping assignments were assigned to the students in CH 301wb. Each concept mapping assignment was completed as a review activity before hour exams. The content of the mapping assignments reflected the material pertinent to the subsequent examination topics. The teaching assistant for the course created the concept list and linking phrases with help from the researcher. Concepts and linking phrases were chosen to limit the possible propositions that students could create. The main concept, or the top-most concept in the hierarchies, was given to the student in all capital letters. This was done as a means to introduce students to concept mapping – a method to get students started on the concept mapping assignment.

Table 5 Concept Mapping Assignment 1

Concepts	**Linking Phrases**
PERIODIC TABLE	arranged by
# of protons	equals
Atomic number	exhibit
Chlorine	has columns called
Diamagnetism	has rows called
Groups	have same
Halogens	members have same
L quantum number	such as
Noble gases	
Paramagnetism	
Periods	

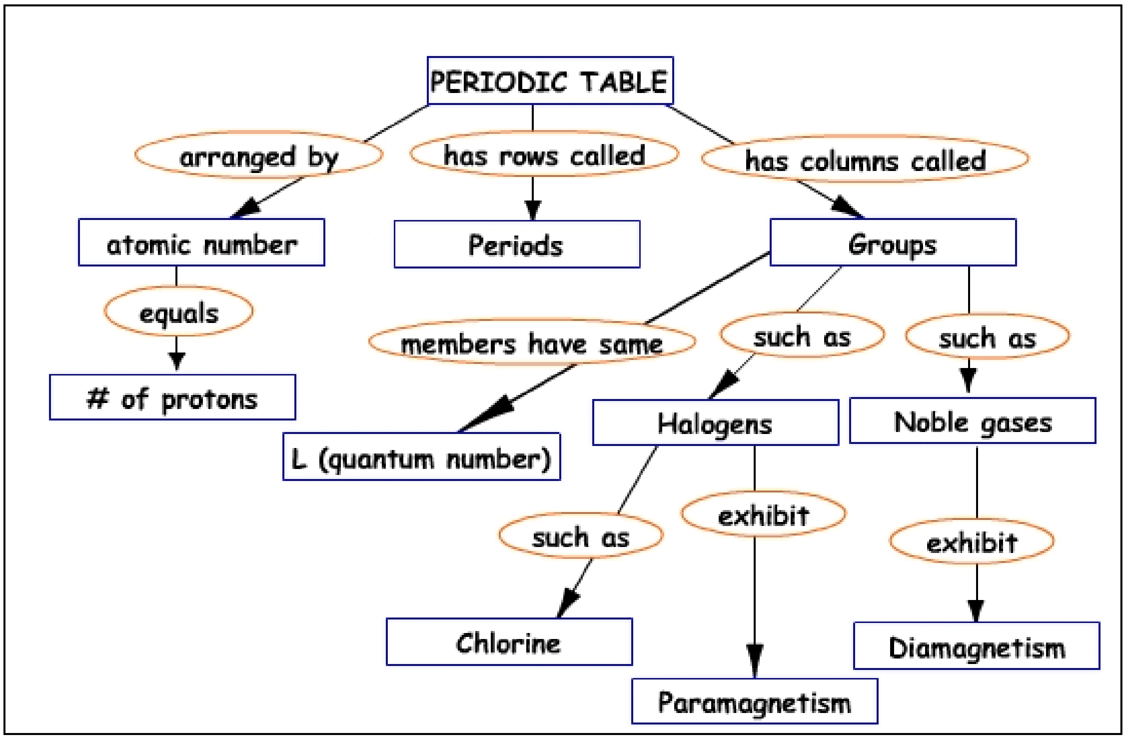


Figure 11 Master map for concept mapping assignment 1.

Table 6 Concept Mapping Assignment 2	
Concepts	**Linking Phrases**
90 degrees	determine
109 degrees	greater repulsion than
BP-BP	has bond angle of
Electronic geometry	has hybridization of
ELECTRONIC REPULSIONS	such as
Linear	
LP-BP	
LP-LP	
Octahedral	
sp	
sp2	
Tetrahedral	
Trigonal planar	

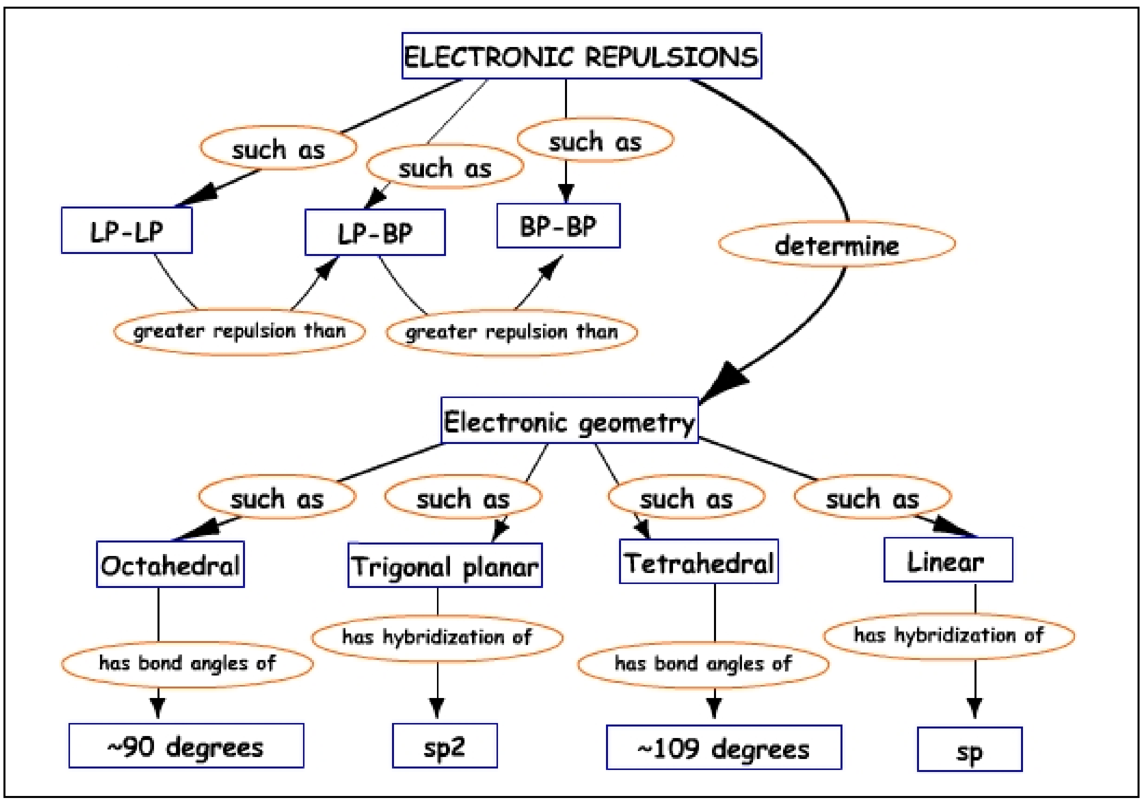


Figure 12 Master map for concept mapping assignment 2.

Table 7 Concept Mapping Assignment 3	
Concepts	**Linking Phrases**
Acid	decreases in
Base	defined as
CHEMICAL REACTIONS	increases in
Insoluble Substance	products include
Metathesis Reaction	reactants include
Neutralization Reaction	such as
Oxidizing Agent	
Oxidation #	
Precipitation Reaction	
Proton Acceptor	
Proton Donor	
Redox Reaction	
Reducing Agent	
Salt	

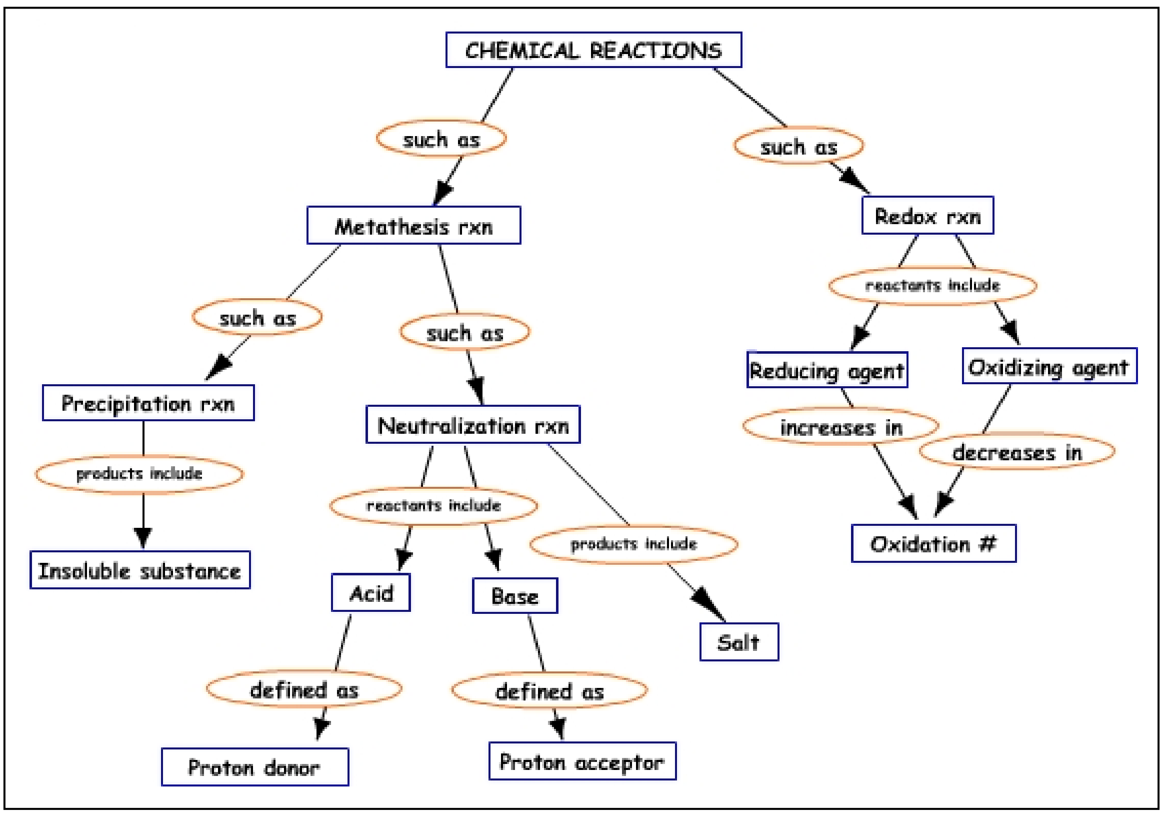


Figure 13 Master map for concept mapping assignment 3.

EXPERIMENTAL DESIGN

Three different experimental approaches were used in this study to investigate the use of CMAT and the usefulness of scoring concept maps by different methods. The data for each of the experiments comes from concept maps drawn by CH 301wb students using the CMAT program. A description of each experiment, the research questions and the testable null hypotheses is given below, followed by a description of the data collection and data analysis.

Experiment 1 – Validating CMAT's automated scoring algorithm

The most important factor in validating the scoring algorithms in the CGI script of the CMAT program is comparing the CMAT generated score of a concept map to the score assigned by a human using the same scoring protocol.

Research Question

Will the scoring algorithm in CMAT score concept maps in a similar fashion as hand grading concept maps?

Null Hypothesis

H_1: A concept map's score from CMAT will be no different than the concept map's score from hand grading.

Experiment 2 – Comparing two relational scoring schemes

The difference in scoring methods is cited as a hindrance for some teachers adopting concept mapping into their classroom [13]. This experiment investigates the differences between concept map scores between two similar scoring schemes. The scoring scheme reported by Francisco and Nakleh is used as a comparison to CMAT scoring [37].

Research questions

Does a using different scoring scheme produce a different score for a concept map? Does the relative difference between concept maps change upon using different scoring schemes?

Null hypotheses

H_2: A concept map score generated by CMAT will be no different than the score generate by the Francisco and Nakleh scheme.

H_3: A rank ordering of concept maps based upon scores generated in CMAT will be no different than the rank ordering based upon scores generated by the Francisco and Nakleh scheme.

Experiment 3 – Comparing structural scoring to relational scoring of concept maps

The information used to calculate the structural complexity index (SCI) is different than that used to calculate the concept map score in CMAT. However, with a limited number of choices in the CMAT interface, significant overlap may exist in the two scoring techniques. Accordingly, it is important to determine what differences, if any, exist in the two scoring schemes. In addition, a verification of the SCI scoring method is needed to address its ability to sort concept maps based upon their structural characteristics.

Research question

Does the SCI approach to scoring concept maps produce a different rank ordering of concept maps than the relational scoring scheme employed by CMAT? What information does the SCI provide towards using the SCI as an assessment of student understanding?

Null hypothesis

H_4: A list of concept maps ordered according to the SCI will be no different than a list of concept maps ordered according the CMAT score.

STUDENT POPULATION

The 32 students enrolled in the web-based general chemistry course (CH 301wb) at The University of Texas at Austin during the spring of 2003 comprised the sample for the study (IRB Protocol # 2002-11-0083). The researcher was not the teaching assistant for the course.

DATA COLLECTION

Students completed 3 concept maps as part of their normal course work. Before the students started their first concept mapping assignment, the teaching assistant offered a help session to show students how to complete a concept map using the CMAT program. Help files were available on-line for the student that described concept mapping and showed students how to use the CMAT program (Appendix A). Students were assigned a unique identification number to enter their work for the course. At the end of the semester, the teaching assistant provided the concept maps and exam answers with the unique identification number and no other information to the researcher.

DATA ANALYSIS

Concept map information, in the form of proposition lists in a text file, was used to recreate concept maps (Figure 14). Each map was then scored by hand using three different scoring schemes: 1) using the CMAT scoring scheme; 2) using the Francisco and Nakleh scoring scheme; and 3) using the SCI scoring scheme.

Experiment 1

Concept map scores generated from the CMAT program and by hand were plotted on a graph. Regression analysis used to determine the degree to which the automated scoring predicted the hand graded score.

Experiment 2

Concept map scores generated from the CMAT program and using the Francisco and Nakleh scoring scheme were plotted on a graph. Regression analysis was used to determine the degree to which the CMAT score predicted the Francisco and Nakleh score. An ANOVA test was performed to determine if the two sets of scores were different from one another; the CMAT score was rescaled to accommodate to the Francisco and Nakleh scoring scheme.

Experiment 3

Concept map scores generated from the CMAT program and using the SCI scoring scheme were plotted on a graph. Regression analysis was used to determine the degree to which the CMAT score predicted the SCI score.

Chapter 4. Results and Discussion

In this chapter the concept maps are analyzed according the methods detailed in the previous chapter. The significance of the results is provided as the results are presented in this chapter.

CONCEPT MAP ANALYSIS

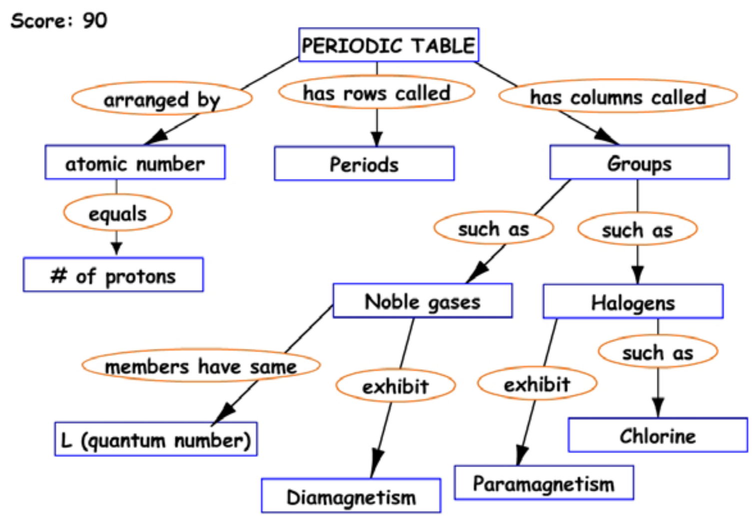
In order to test the hypothesis in this study, concept maps submitted by the student participants in the study had to be first transformed into a sensible format to replicate human grading (See Appendix B, C, and D). Five examples of the transformation process and scoring procedures are shown in Figures 14 – 19. The concept maps submitted from each were recorded by the CGI script of the CMAT program into a text file (Figure 14a). The list of propositions was converted into a graphical concept map (Figure 14b). This concept map was then scored by the researcher using three different scoring methods; the necessary information for the different scoring methods (Figure 14c) was recorded for each map. Finally, the concept map score, using the scoring methods, was recorded for each concept map (Figure 14d).

(a)

Student 27 90

1. PERIODIC TABLE, arranged by, Atomic number
2. PERIODIC TABLE, has columns called, Groups
3. PERIODIC TABLE, has rows called, Periods
4. Atomic number, equals,# of protons
5. Halogens, exhibit, paramagnetism
6. noble gases, exhibit, diamagnetism
7. noble gases, members have same, L (quantum number)
8. Halogens, such as, Chlorine
9. Groups, such as, Halogens
10. Groups, such as, noble gases

(b)



(c)

Concept Map Scoring Criteria

Possible propositions	- 10
Propositions made	- 10
Correct propositions	- 9
Wrong propositions	- 1
Non-informative props.	- 0
Branching points	- 4
Chains	- 6
Average chain length	- 2.5

(d)

Scoring Method	Score
CMAT	90

By hand

$$50 + 50 \times (9 - 1)/10 = 90$$

Francisco and Nakleh

$$(9 - (1 + 0))/10 \times 100 = 80$$

Structural Complexity Index

$$(2.5 \times 10) + (4 \times 6) = 49$$

The data analysis for Student 27 concept mapping assignment 1 is shown here. The concept recorded by the CGI script of the CMAT program (a) is transformed into a concept map (b). The propositions of the concept map are coded into the categories used in the different scoring schemes (c). The different scores for the concept map are calculated (d).

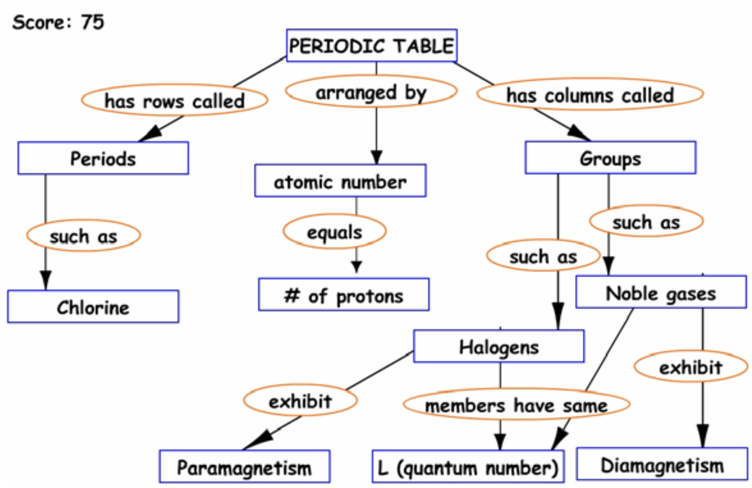
Figure 14 Concept map analysis #1

(a)

Student 15 75

1. PERIODIC TABLE, has columns called, Groups
2. PERIODIC TABLE, has rows called, Periods
3. PERIODIC TABLE, arranged by, Atomic number
4. Atomic number, equals, # of protons
5. Groups, such as, noble gases
6. Groups, such as, Halogens
7. Periods, such as, Chlorine
8. Halogens, members have same, L (quantum number)
9. Halogens, exhibit, paramagnetism
10. noble gases, members have same, L (quantum number)
11. noble gases, exhibit, diamagnetism

(b)



(c)

Concept Map Scoring Criteria

Possible propositions	- 10
Propositions made	- 11
Correct propositions	- 8
Wrong propositions	- 1
Non-informative props.	- 2
Branching points	- 4
Chains	- 6
Average chain length	- 2.67

(d)

Scoring Method	Score
CMAT	75

By hand

$$50 + 50 \times (8 - 3)/10 = 75$$

Francisco and Nakleh

$$(8 - (1 + 2))/11 \times 100 = 45$$

Structural Complexity Index

$$(2.67 \times 11) + (4 \times 6) = 53.4$$

The data analysis for Student 15 concept mapping assignment 1 is shown here. The concept recorded by the CGI script of the CMAT program (a) is transformed into a concept map (b). The propositions of the concept map are coded into the categories used in the different scoring schemes (c). The different scores for the concept map are calculated (d).

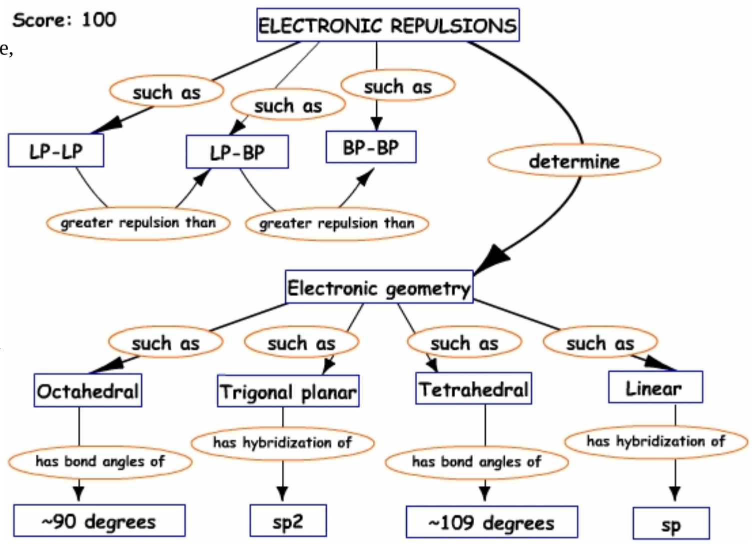
Figure 15 Concept map analysis #2

(a)

Student 27 100

1. ELECTRONIC REPULSIONS, determine, Electronic geometry
2. LP-LP, greater repulsion than, LP-BP
3. LP-BP, greater repulsion than, BP-BP
4. ELECTRONIC REPULSIONS, such as, LP-LP
5. ELECTRONIC REPULSIONS, such as, BP-BP
6. ELECTRONIC REPULSIONS, such as, LP-BP
7. Electronic geometry, such as, Linear
8. Electronic geometry, such as, Octahedral
9. Electronic geometry, such as, Tetrahedral
10. Electronic geometry, such as, Trigonal planar
11. Linear, has hybridization of, sp
12. Octahedral, has bond angles of, ~90 degrees
13. Tetrahedral, has bond angles of, ~109 degrees
14. Trigonal planar, has hybridization of, sp2

(b)



(c)

Concept Map Scoring Criteria

Possible propositions	- 14
Propositions made	- 14
Correct propositions	- 14
Wrong propositions	- 0
Non-informative props.	- 0
Branching points	- 2
Chains	- 7
Average chain length	- 2.57

(d)

Scoring Method	Score
CMAT	100

By hand

$$50 + 50 \times (14)/14 = 100$$

Francisco and Nakleh

$$(14 - (0))/14 \times 100 = 100$$

Structural Complexity Index

$$(2.57 \times 14) + (2 \times 7) = 50.0$$

The data analysis for Student 27 concept mapping assignment 2 is shown here. The concept recorded by the CGI script of the CMAT program (a) is transformed into a concept map (b). The propositions of the concept map are coded into the categories used in the different scoring schemes (c). The different scores for the concept map are calculated (d).

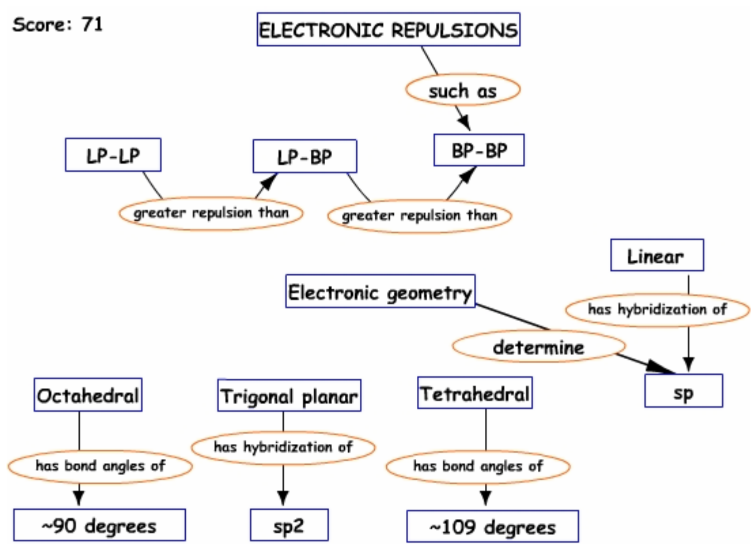
Figure 16 Concept map analysis #3.

(a)

Student 15 71

1. Octahedral, has bond angles of, ~90 degrees
2. Tetrahedral, has bond angles of, ~109 degrees
3. Linear, has hybridization of,sp
4. Trigonal planar, has hybridization of, sp2
5. Electronic geometry, determine, sp
6. ELECTRONIC REPULSIONS, such as, BP-BP
7. LP-LP, greater repulsion than, LP-BP
8. LP-BP, greater repulsion than, BP-BP

(b)



(c)

Concept Map Scoring Criteria

Possible propositions	- 14
Propositions made	- 12
Correct propositions	- 7
Wrong propositions	- 1
Non-informative props.	- 4
Branching points	- 0
Chains	- 7
Average chain length	- 1.14

(d)

Scoring Method	Score
CMAT	71

By hand

$$50 + 50 \times (7 - 5)/14 = 57$$

Francisco and Nakleh

$$(7 - (1 + 4))/12 \times 100 = 17$$

Structural Complexity Index

$$(1.14 \times 12) + (0 \times 8) = 13.7$$

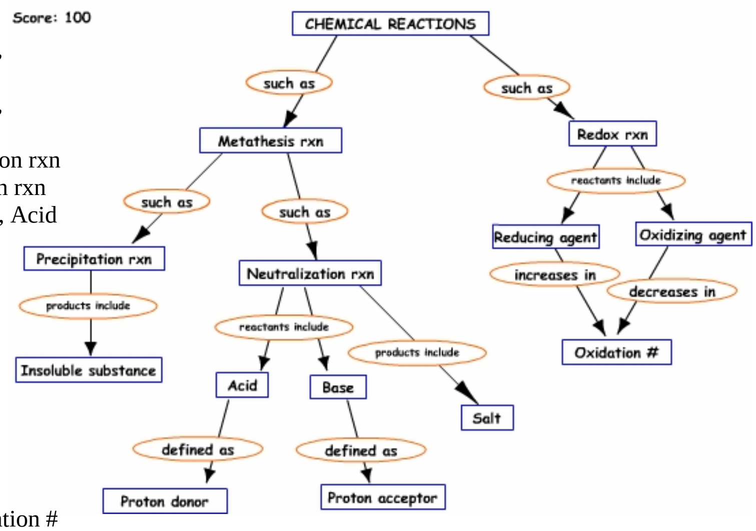
The data analysis for Student 15 concept mapping assignment 2 is shown here. The concept recorded by the CGI script of the CMAT program (a) is transformed into a concept map (b). The propositions of the concept map are coded into the categories used in the different scoring schemes (c). The different scores for the concept map are calculated (d).

Figure 17 Concept map analysis #4.

(a) (b)

Student 27 100

1. CHEMICAL REACTIONS, such as, Redox rxn
2. CHEMICAL REACTIONS, such as, Metathesis rxn
3. Metathesis rxn, such as, Neutralization rxn
4. Metathesis rxn, such as, Precipitation rxn
5. Neutralization rxn, reactants include, Acid
6. Neutralization rxn, reactants include, Base
7. Neutralization rxn, products include, Salt
8. Precipitation rxn, products include, Insoluble substance
9. Redox rxn, reactants include, Oxidizing agent
10. Redox rxn, reactants include, Reducing agent
11. Oxidizing agent, decreases in, Oxidation #
12. Reducing agent, increases in, Oxidation #
13. Acid, defined as, Proton donor
14. Base, defined as, Proton acceptor



(c)

Concept Map Scoring Criteria

Possible propositions	- 14
Propositions made	- 14
Correct propositions	- 14
Wrong propositions	- 0
Non-informative props.	- 0
Branching points	- 4
Chains	- 6
Average chain length	- 3.33

(d)

Scoring Method	Score
CMAT	100

By hand

$$50 + 50 \times (14 - 0)/14 = 100$$

Francisco and Nakleh

$$(14 - (0 + 0))/14 \times 100 = 100$$

Structural Complexity Index

$$(3.33 \times 14) + (4 \times 6) = 70.6$$

The data analysis for Student 27 concept mapping assignment 3 is shown here. The concept recorded by the CGI script of the CMAT program (a) is transformed into a concept map (b). The propositions of the concept map are coded into the categories used in the different scoring schemes (c). The different scores for the concept map are calculated (d).

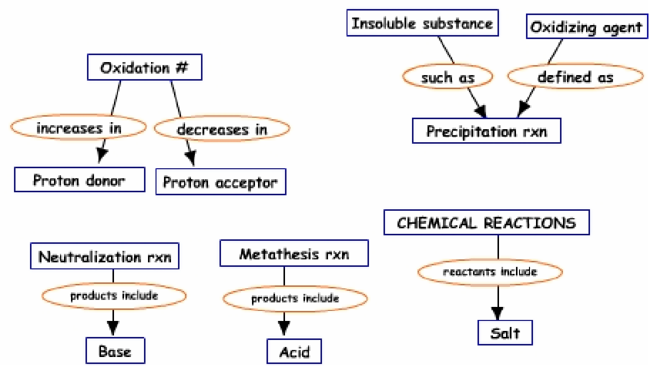
Figure 18 Concept map analysis #5.

(a)

Student 15 29

1. Oxidation #, increases in, Proton donor
2. Oxidation #,decreases in, Proton acceptor
3. Insoluble substance, such as, Precipitation rxn
4. CHEMICAL REACTIONS, reactants include, Salt
5. Oxidizing agent, defined as, Precipitation rxn
6. Neutralization rxn, products include, Base
7. Metathesis rxn, products include, Acid

(b)



(c)

<u>Concept Map Scoring Criteria</u>

Possible propositions	- 14
Propositions made	- 7
Correct propositions	- 0
Wrong propositions	- 7
Non-informative props.	- 0
Branching points	- 1
Chains	- 7
Average chain length	- 1.00

(d)

Scoring Method	Score
CMAT	29

By hand

$$50 + 50 \times (0 - 7)/14 = 25$$

Francisco and Nakleh

$$(0 - (0 + 7))/7 \times 100 = -100$$

Structural Complexity Index

$$(1.00 \times 7) + (1 \times 7) = 14$$

The data analysis for Student 15 concept mapping assignment 3 is shown here. The concept recorded by the CGI script of the CMAT program (a) is transformed into a concept map (b). The propositions of the concept map are coded into the categories used in the different scoring schemes (c). The different scores for the concept map are calculated (d).

Figure 19 Concept map analysis #6.

In the study, students enrolled in the CH 301wb class drew concept maps using the CMAT program; the level of participation is shown in Table 8.

Table 8 Participation rate for each concept mapping activity.

Concept Mapping Activity	1	2	3
Number of students submitting a concept map	25	21	22

For each concept map drawn in the CMAT program, an electronic score was generated and all of the propositions recorded in a file. From the recorded propositions, the concept map was re-created and scored by hand, by the researcher. The concept map score was determined by examining each concept map and classifying each proposition created as correct, incorrect and/or non–informative. The non–informative category was used if independent propositions were created outside of the main hierarchy of links. The data for each concept map drawn in the three concept mapping activities is shown in Tables 9 – 11.

Table 9 Concept map analysis results from Concept Mapping Assignment 1.

Propositions were counted as correct, wrong, or non–informative; there were 10 possible correct propositions as determined by the master map. Human score calculated using Equation 1, CMAT score was generated by the CMAT program. For students who did not submit a concept map, N/A is entered.

	# of links classified as			Concept map score via	
Student	**Correct**	**Non-informative**	**Wrong**	**Human**	**CMAT (electronic)**
Student 1	10	0	0	100	95
Student 2	7	2	2	65	65
Student 3	5	0	0	75	75
Student 4	7	0	0	85	85
Student 5	5	0	0	75	75
Student 6	7	0	0	85	85
Student 7	N/A	N/A	N/A	N/A	N/A
Student 8	7	0	0	85	75
Student 9	5	0	0	75	75
Student 10	5	2	1	60	65
Student 11	5	0	0	75	75
Student 12	4	2	1	55	65
Student 13	5	0	0	75	75
Student 14	7	2	0	75	70
Student 15	8	2	1	75	75
Student 16	N/A	N/A	N/A	N/A	N/A
Student 17	5	0	4	55	55
Student 18	N/A	N/A	N/A	N/A	N/A
Student 19	8	1	0	85	90
Student 20	N/A	N/A	N/A	N/A	N/A
Student 21	4	0	1	65	65
Student 22	N/A	N/A	N/A	N/A	N/A
Student 23	9	1	0	90	95
Student 24	2	0	7	25	20
Student 25	7	0	2	75	75
Student 26	N/A	N/A	N/A	N/A	N/A
Student 27	9	1	0	90	90
Student 28	8	1	0	85	90
Student 29	N/A	N/A	N/A	N/A	N/A
Student 30	7	2	0	75	85
Student 31	7	0	3	70	70
Student 32	7	0	0	85	85

Table 10 Concept map analysis results from Concept Mapping Assignment 2.

Propositions were counted as correct, wrong, or non–informative; there were 14 possible correct propositions as determined by the master map. Human score calculated using Equation 1, CMAT score was generated by the CMAT program. For students who did not submit a concept map, N/A is entered.

	# of links classified as			Concept map score via	
Student ID	**Correct**	**Non-informative**	**Wrong**	**Human**	**CMAT (electronic)**
Student 1	13	0	1	93	93
Student 2	N/A	N/A	N/A	N/A	N/A
Student 3	6	0	0	71	71
Student 4	9	0	0	82	82
Student 5	12	0	0	93	93
Student 6	9	0	0	82	82
Student 7	N/A	N/A	N/A	N/A	N/A
Student 8	9	0	0	82	82
Student 9	5	0	0	68	68
Student 10	14	0	0	100	100
Student 11	N/A	N/A	N/A	N/A	N/A
Student 12	9	0	5	64	64
Student 13	N/A	N/A	N/A	N/A	N/A
Student 14	12	0	0	93	93
Student 15	7	4	1	57	71
Student 16	5	0	0	68	68
Student 17	N/A	N/A	N/A	N/A	N/A
Student 18	N/A	N/A	N/A	N/A	N/A
Student 19	14	0	0	100	100
Student 20	N/A	N/A	N/A	N/A	N/A
Student 21	5	0	0	68	68
Student 22	N/A	N/A	N/A	N/A	N/A
Student 23	8	0	2	71	75
Student 24	13	0	1	93	93
Student 25	8	1	2	68	68
Student 26	N/A	N/A	N/A	N/A	N/A
Student 27	14	0	0	100	100
Student 28	13	3	0	86	86
Student 29	N/A	N/A	N/A	N/A	N/A
Student 30	12	0	0	93	93
Student 31	7	7	1	46	61
Student 32	N/A	N/A	N/A	N/A	N/A

Table 11 Concept map analysis results from Concept Mapping Assignment 3.

Propositions were counted as correct, wrong, or non–informative; there were 14 possible correct propositions as determined by the master map. Human score calculated using Equation 1, CMAT score was generated by the CMAT program. For students who did not submit a concept map, N/A is entered.

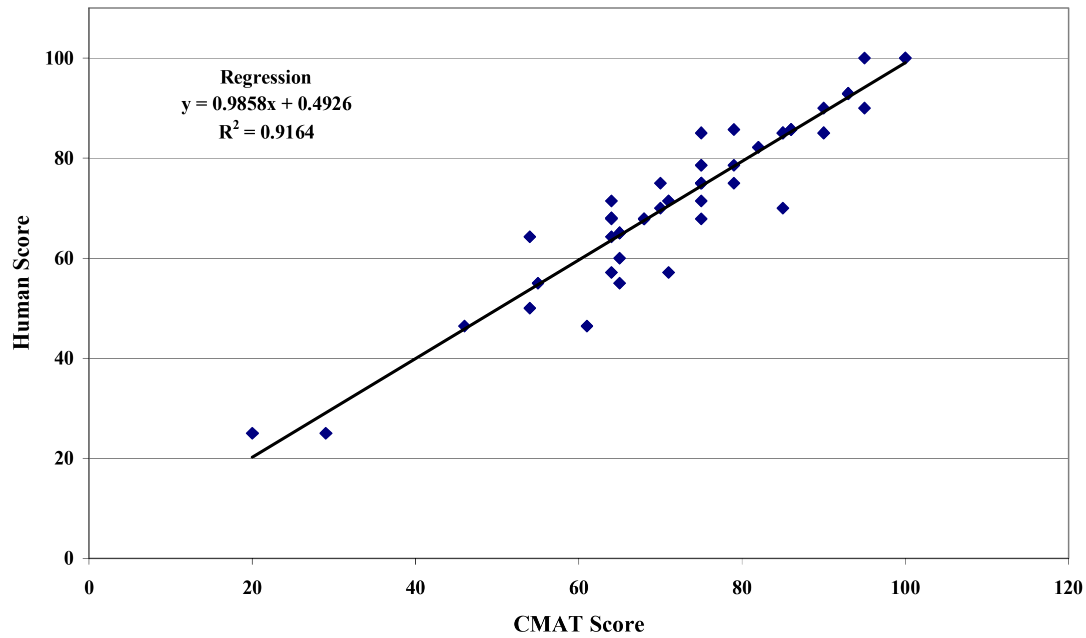
	# of links classified as			Concept map score via	
Student ID	**Correct**	**Non-informative**	**Wrong**	**Human**	**CMAT (electronic)**
Student 1	12	1	1	86	86
Student 2	12	3	2	75	79
Student 3	8	0	0	79	75
Student 4	12	0	2	86	79
Student 5	7	0	2	68	64
Student 6	N/A	N/A	N/A	N/A	N/A
Student 7	10	3	2	68	75
Student 8	6	0	2	64	54
Student 9	5	3	0	57	64
Student 10	12	0	2	86	86
Student 11	N/A	N/A	N/A	N/A	N/A
Student 12	11	3	2	71	71
Student 13	9	0	1	79	75
Student 14	N/A	N/A	N/A	N/A	N/A
Student 15	0	0	7	25	29
Student 16	5	0	0	68	64
Student 17	10	0	3	75	75
Student 18	N/A	N/A	N/A	N/A	N/A
Student 19	12	0	2	86	86
Student 20	N/A	N/A	N/A	N/A	N/A
Student 21	2	1	2	46	46
Student 22	N/A	N/A	N/A	N/A	N/A
Student 23	12	2	2	79	79
Student 24	6	1	5	50	54
Student 25	10	2	2	71	64
Student 26	N/A	N/A	N/A	N/A	N/A
Student 27	14	0	0	100	100
Student 28	14	2	0	93	93
Student 29	N/A	N/A	N/A	N/A	N/A
Student 30	12	2	0	86	86
Student 31	N/A	N/A	N/A	N/A	N/A
Student 32	N/A	N/A	N/A	N/A	N/A

EXPERIMENT 1

Test of Null Hypothesis 1

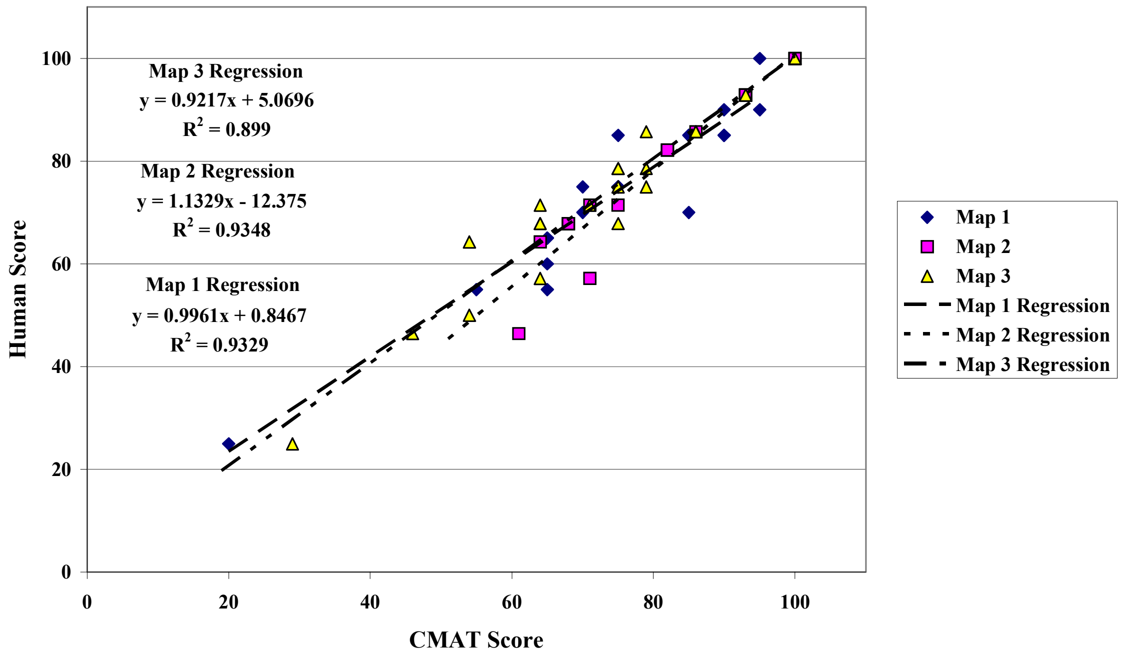
H_1: A concept map's score from CMAT will be the same as than the concept map's score from hand grading.

To answer the question whether the scoring of concept maps in the CMAT program is comparable to hand grading, concept maps scores were compared. A plot of each map's electronic and human score is shown in Figure 20. The same data is shown according to each of the particular concept mapping assignment in Figure 21. The correlation between human scores and electronic scores generated in CMAT is done by regression analysis.



The electronic score (CMAT Score) and corresponding human score for each co ncept map drawn in the study. A number of data points on the graph represent more than one concept maps, for example, 6 concepts maps have a 93 CMAT score and 93 Human score.

Figure 20 Plot of concept map scores generated by hand versus CMAT.



The electronic score (CMAT Score) and corresponding human score for each concept map in the different concept mapping assignments – abbreviated Map 1, Map 2, and Map 3..

Figure 21 Plot of concept map scores generated by hand versus CMAT, separated by concept mapping assignment.

The regression statistics for Figure 20 and Figure 21 are summarized in Table 11. The R^2 value for the correlation of electronic scores generated by CMAT and human scores when considering all concept maps in the study is 0.9234 with a slope of 0.99 +/– 0.035. The slope of the regression line indicates that, on average, the CMAT score will be 99% +/– 3.5% of the score assigned by a human. Similar results are found when considering each concept mapping assignment independently. The lowest correlation occurs in concept mapping assignment 1 which has the fewest number of concepts and possible propositions (10) to make. Concept mapping assignments 2 and 3 both have 14 possible propositions and very similar correlation statistics to human scoring. This would indicate the number of possible propositions in a concept mapping activity influences the degree to which electronic scoring in CMAT correlates to human scores of the same concept maps.

Table 12 Regression statistics for the comparison of CMAT scores and human scores of concept maps.

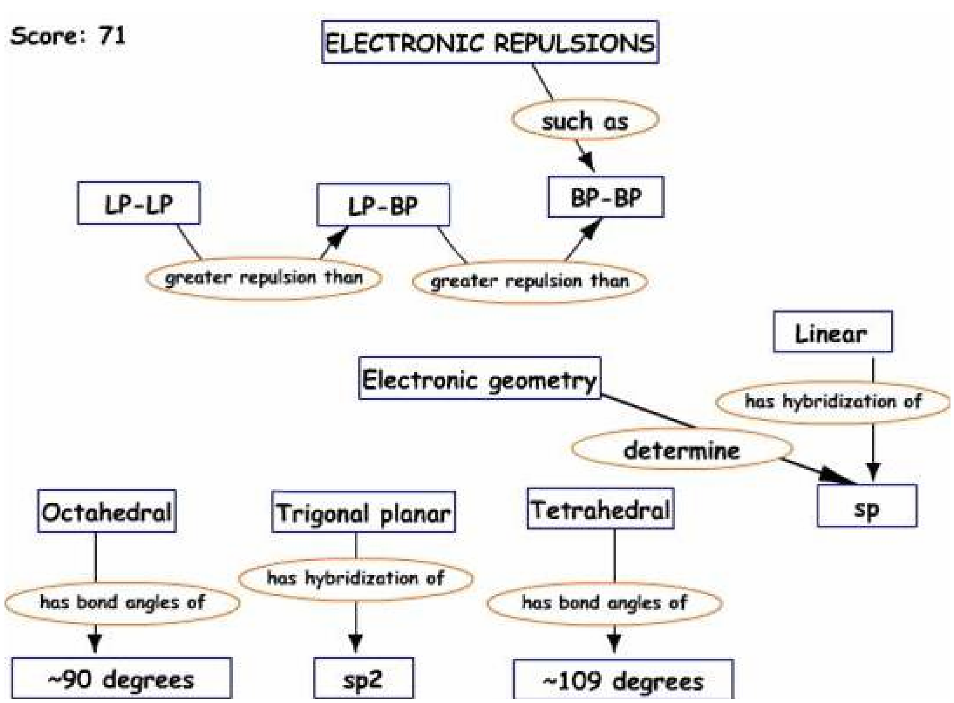
	R^2	Slope	95% Confidence Interval
All mapping assignments combined			
All Maps	**0.9234**	**0.99 ± 0.035**	**0.92 − 1.06**
Individual mapping assignments			
Map 1	**0.9191**	**0.93 ± 0.058**	**0.79 − 1.05**
Map 2	**0.9348**	**1.13 ± 0.069**	**0.99 − 1.28**
Map 3	**0.9329**	**1.00 ± 0.060**	**0.81 − 1.12**

Based upon the correlation between hand scoring concept maps and electronically scored concept maps summarized in Table 11, the Null Hypothesis, H_1, is accepted. A concept map's score obtained from CMAT is no different than the concept map's score obtained from hand grading.

Other Observations from Experiment 1

Why electronic scoring and human scoring is not a "perfect" match?

The electronic score of a concept map generated by CMAT is a ratio of correct propositions to the total possible correct propositions in a concept mapping assignment. The correct propositions are encoded by the instructor by having the instructor creating a master concept map. One assumption that is made in scoring concept maps is that the propositions created by the student create a contiguous hierarchy. However, as demonstrated by Student 15 in concept mapping assignment 2 (shown in Figure 22), student's do not always follow directions with 100% accuracy. Student 15's concept map in shown in Figure 22 contains many correct propositions when considered independently. In the CMAT electronic scoring algorithm, Student 15 only made 1 incorrect proposition and made 7 correct propositions for a score of 71. However, when a human scored the same concept map, 4 of the propositions were considered correct, but non–informative, because were not contained in a contiguous hierarchy. Therefore, one of the drawbacks to electronic scoring of concept maps, as exemplified in the CMAT program, is the difficulty in dealing with non–standard input, and suggests that students need to be carefully trained in the generation of concept maps, whether through “paper–and–pencil” or electronic means.



Concept map drawn by Student 15 for concept mapping assignment 2. Example of a concept map where all the propositions were not included in a contiguous hierarchy.

Figure 22 Example of a poorly drawn concept map.

Does the number of concepts included in the concept mapping assignment affect the correlation?

While this study was not designed to address this question, there appears to be some effect in the correlation between hand scoring maps and electronic scoring in the CMAT program when the number of concepts used is small. Concept mapping assignment 1 has a lower correlation coefficient than the other two concept mapping assignments. A rationalization can be made that with fewer concepts, a single difference in scoring will have more of an affect. As concept maps are being proposed as an assessment tool, the number of concepts included in a concept mapping assignment to make it an effective assessment should be studied further.

EXPERIMENT 2 – COMPARING TWO RELATIONAL SCORING SCHEMES

Test of Null hypotheses

H_2: A concept map score generated by CMAT will be the same as than the score generated by the Francisco and Nakleh scheme.

H_3: A rank ordering of concept maps based upon scores generated in CMAT will be the same as the rank ordering based upon scores generated by the Francisco and Nakleh scheme.

The concept maps drawn and scored in CMAT were scored according the scoring scheme reported by Francisco and Nakleh et al. This scoring scheme is very similar to the one used in CMAT with two exceptions: (1) only the total propositions used by the student are considered in the Nakleh method, whereas in CMAT, the total possible proposition are considered; (2) scores in CMAT are normalized to avoid negative scores, whereas in the Nakleh method, scores can range from –100% to +100%. Both scoring methods can be considered *relational* in nature because they rely on the accuracy of the propositions for their score. The scores for each of the concept maps drawn in CMAT were assigned a score according to the Francisco and Nakleh method (Table 12), see Figures 14–19 for and example calculations.

Table 13 Concept maps scores generated by CMAT and the Francisco and Nakleh method. For students who did not submit a concept map, N/A is entered.

	Concept Map 1		Concept Map 2		Concept Map 3	
Student	**CMAT (electronic)**	**Nakleh method**	**CMAT (electronic)**	**Nakleh method**	**CMAT (electronic)**	**Nakleh method**
Student 1	95	100	93	86	86	71
Student 2	65	27	N/A	N/A	79	41
Student 3	75	100	71	100	75	100
Student 4	85	100	82	100	79	71
Student 5	75	100	93	100	64	56
Student 6	85	100	82	100	N/A	N/A
Student 7	N/A	N/A	N/A	N/A	75	33
Student 8	75	100	82	100	54	50
Student 9	75	100	68	100	64	25
Student 10	65	25	100	100	86	71
Student 11	75	100	N/A	N/A	N/A	N/A
Student 12	65	14	64	29	71	38
Student 13	75	100	N/A	N/A	75	80
Student 14	70	56	93	100	N/A	N/A
Student 15	75	45	71	17	29	-100
Student 16	N/A	N/A	68	100	64	100
Student 17	55	11	N/A	N/A	75	54
Student 18	N/A	N/A	N/A	N/A	N/A	N/A
Student 19	90	78	100	100	86	71
Student 20	N/A	N/A	N/A	N/A	N/A	N/A
Student 21	65	60	68	100	46	-20
Student 22	N/A	N/A	N/A	N/A	N/A	N/A
Student 23	95	80	75	60	79	50
Student 24	20	-56	93	86	54	0
Student 25	75	56	68	45	64	43
Student 26	N/A	N/A	N/A	N/A	N/A	N/A
Student 27	90	80	100	100	100	100
Student 28	90	78	86	63	93	75
Student 29	N/A	N/A	N/A	N/A	N/A	N/A
Student 30	85	56	93	100	86	71
Student 31	70	40	61	-7	N/A	N/A
Student 32	85	100	N/A	N/A	N/A	N/A

In Figure 23 the concept map scores generated by CMAT and the Nakleh scoring method are plotted for each concept map drawn in the study. The regression line drawn (data shown in Table 13) shows the two methods have a weak correlation ($R^2 = 0.5847$).

The slope of the line between the two scoring methods, however, is exactly as would be expected, that is a value of 2, because the range of possible scores for CMAT (0 to 100) is half that of the Nakleh scoring method (–100 to 100).

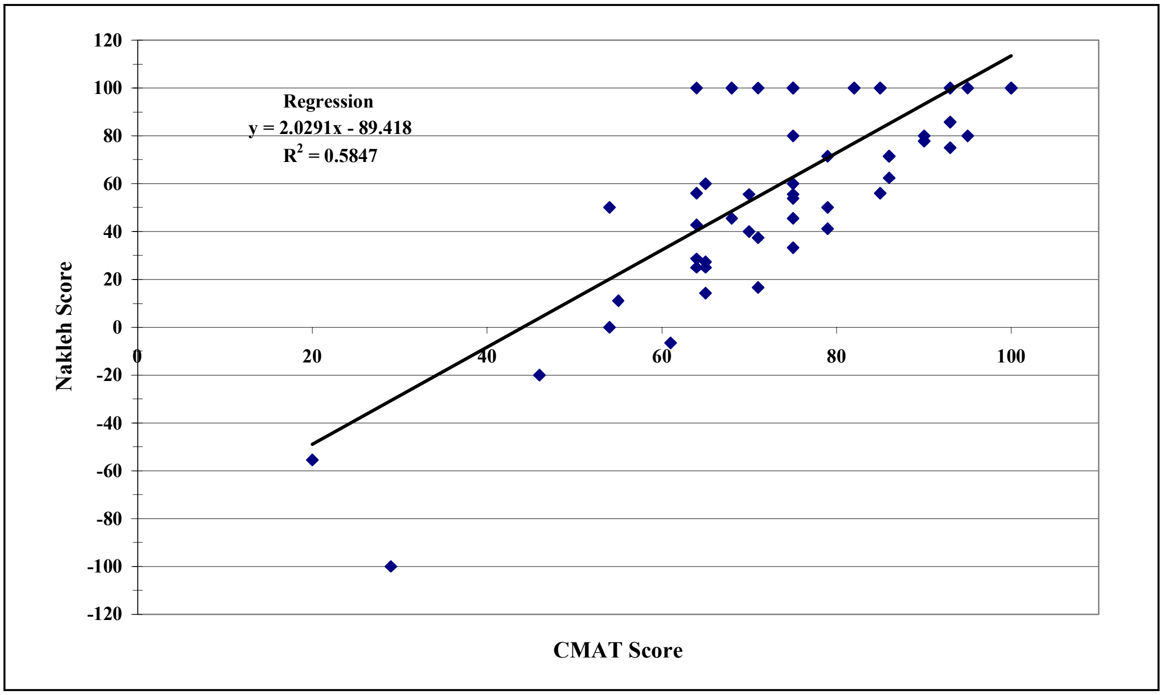


Figure 23 Plot of all concept map scores generated by Francisco and Nakleh method versus CMAT.

Table 14 Regression data for the plot of concept map scores generated by CMAT and the Nakleh scoring method.

	R^2	Slope	95% Confidence Interval
All Maps	**0.5847**	**2.0 ± 0.21**	**1.6 − 2.4**

For the Null Hypothesis, H_2, that a concept map score generated by CMAT will be the same as the score generated by the Francisco and Nakleh scheme can be accepted based upon the slope of the regression line. When averaged over a large sample size, the

Francisco and Nakleh scoring method and CMAT scoring algorithm, both relational scoring methods, will score concept maps similarly.

For the Null Hypothesis, H_3, that a rank ordering of concept maps based upon scores generated in CMAT will be the same as the rank ordering based upon scores generated by the Francisco and Nakleh scheme must be rejected based upon the correlation coefficient of regression analysis in Table 13. An R^2 value of 0.5847 means that knowing the CMAT score will reduce the variance in predicting the concept map scoring using the Francisco and Nakleh method by 58%. This poor fit indicates that students who scored high on one method did not necessarily score high on the other method.

The weak correlation between scores from the different scoring schemes is disconcerting and leads to other questions: "Why do the scores differ?" or "Is one score *better* than the other?"

Why do concept map scores differ?

In the Nakleh method, only propositions created by the student are considered. This process differs from the CMAT approach to scoring where the score is based upon the propositions the students *should* create with the concepts provided in the concept mapping assignment. In the present comparison, a student who correctly uses just a few of the concepts is scored highly by the Nakleh method but is not scored very highly in the CMAT scheme; clearly, the difference in approaches in philosophical. If students used all of the concepts to create their concept maps, we would expect that the scores generated by the Nakleh scoring method and CMAT would converge.

Is one score *better* than the other?

Both scoring schemes accomplish the stated purpose of their creators. The Francisco and Nakleh scoring scheme details how well a student creates a concept map with the concepts they include. The CMAT scoring scheme details how well a student can create a concept map using terms the instructor thinks are pertinent to the concept mapping assignment. The Francisco and Nakleh scheme is "passive" in the sense that it does not penalize the student for not using all the possible propositions, in contrast to the CMAT approach. The difference can be explained with the instructor's expectations of student behaviors. Therefore, the question of *better* addresses the purpose of the concept mapping assignment. If the purpose of the concept mapping assignment is to determine the quality of a student's understanding, than the Nakleh scoring method would seem appropriate. If the purpose of the concept mapping assignment is determine the quality of a student's understanding over a given range of topics then the CMAT scoring approach would be appropriate.

The last factor to consider in determining *better* is the scoring environment – electronic or in–person. When scoring concept maps in–person, decisions about quality based upon the number of concepts used can be factored into the scoring algorithm, each map can be considered independently. However, the human scorer must maintain a consistency throughout the scoring of all concept maps. This condition is similar to hand grading of textual responses; consistency across all student answers becomes a problem. The CMAT approach is, if anything, consistent. When scoring concept maps electronically, as in CMAT, the scoring scheme cannot be adjusted for each concept maps and needs to be robust in accounting for different numbers of concept used. Therefore, for the purposes of creating an electronic concept mapping system that can evaluate a student's concept map, the scoring scheme employed by CMAT is better than the Nakleh scoring method.

EXPERIMENT 3 – COMPARING STRUCTURAL SCORING TO RELATIONAL SCORING OF CONCEPT MAPS

During the course of this study, a new metric for evaluating concept maps was created, the Structural Complexity Index (SCI). The information used to calculate the SCI is different than that used to calculate the concept map score in CMAT (see Figures 14–19). Whether or not concepts scored within the CMAT program are equivalently evaluated using the SCI is one of the questions addressed by this study.

Test of the Null Hypothesis

H_4: A list of concept maps ordered according to the SCI will be no different than a list of concept maps ordered according the CMAT score.

The concept maps created and scored with the CMAT program were scored according to the SCI (equation 2, Chapter 3), see Table 14. A plot of the CMAT score and corresponding SCI (Figure 24) was used to determine the relationship between a map's SCI and CMAT score. The regression analysis for the plot in Figure 24 is shown in Table 15.

Table 15 Concept maps scores according to CMAT program and the Structural Complexity Index. For students who did not submit a concept map, N/A is entered.

	Concept Mapping Activity 1		**Concept Mapping Activity 2**		**Concept Mapping Activity 3**	
Student	**CMAT (electronic)**	**Structural Complexity Index**	**CMAT (electronic)**	**Structural Complexity Index**	**CMAT (electronic)**	**Structural Complexity Index**
Student 1	95	41	93	50	86	71
Student 2	65	16	N/A	N/A	79	52
Student 3	75	11	71	19	75	32
Student 4	85	22	82	31	79	60
Student 5	75	16	93	42	64	37
Student 6	85	22	82	31	N/A	N/A
Student 7	N/A	N/A	N/A	N/A	75	41
Student 8	75	22	82	31	54	32
Student 9	75	11	68	17	64	8
Student 10	65	28	100	50	86	71
Student 11	75	11	N/A	N/A	N/A	N/A
Student 12	65	26	64	70	71	33
Student 13	75	16	N/A	N/A	75	41
Student 14	70	17	93	46	N/A	N/A
Student 15	75	51	71	3	29	8
Student 16	N/A	N/A	68	17	64	16
Student 17	55	33	N/A	N/A	75	74
Student 18	N/A	N/A	N/A	N/A	N/A	N/A
Student 19	90	31	100	50	86	71
Student 20	N/A	N/A	N/A	N/A	N/A	N/A
Student 21	65	11	68	17	46	10
Student 22	N/A	N/A	N/A	N/A	N/A	N/A
Student 23	95	34	75	53	79	67
Student 24	20	47	93	48	54	47
Student 25	75	30	68	29	64	48
Student 26	N/A	N/A	N/A	N/A	N/A	N/A
Student 27	90	49	100	50	100	71
Student 28	90	24	86	154	93	55
Student 29	N/A	N/A	N/A	N/A	N/A	N/A
Student 30	85	33	93	40	86	47
Student 31	70	41	61	18	N/A	N/A
Student 32	85	22	N/A	N/A	N/A	N/A

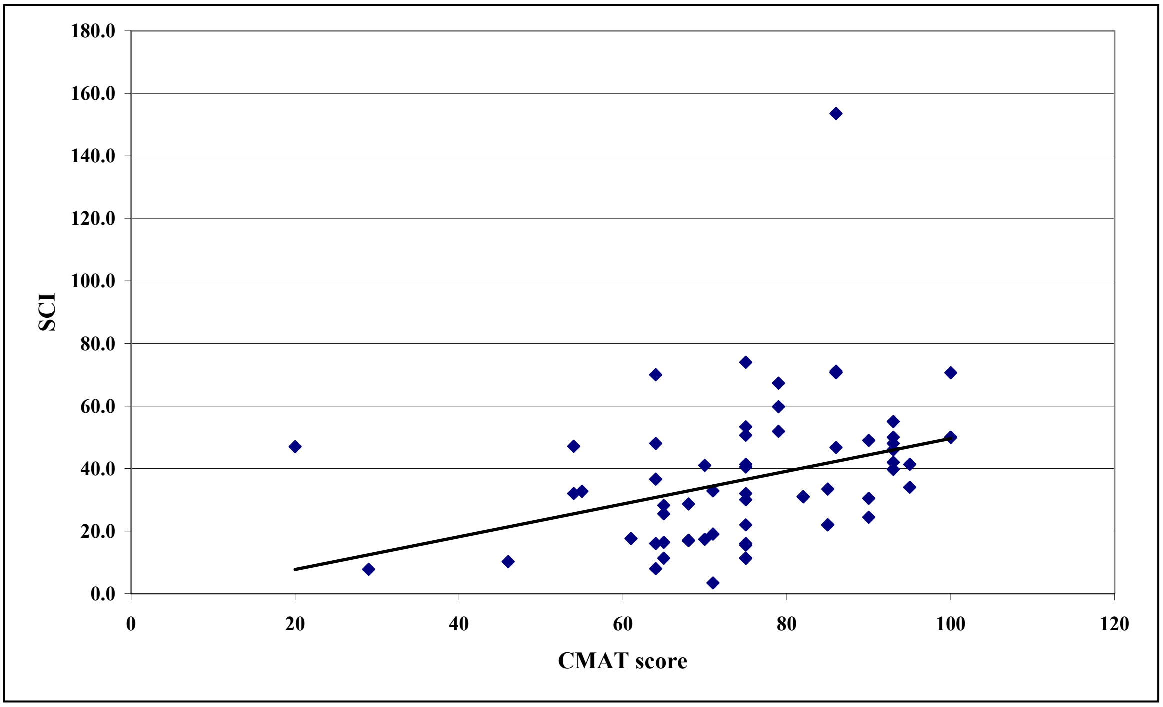
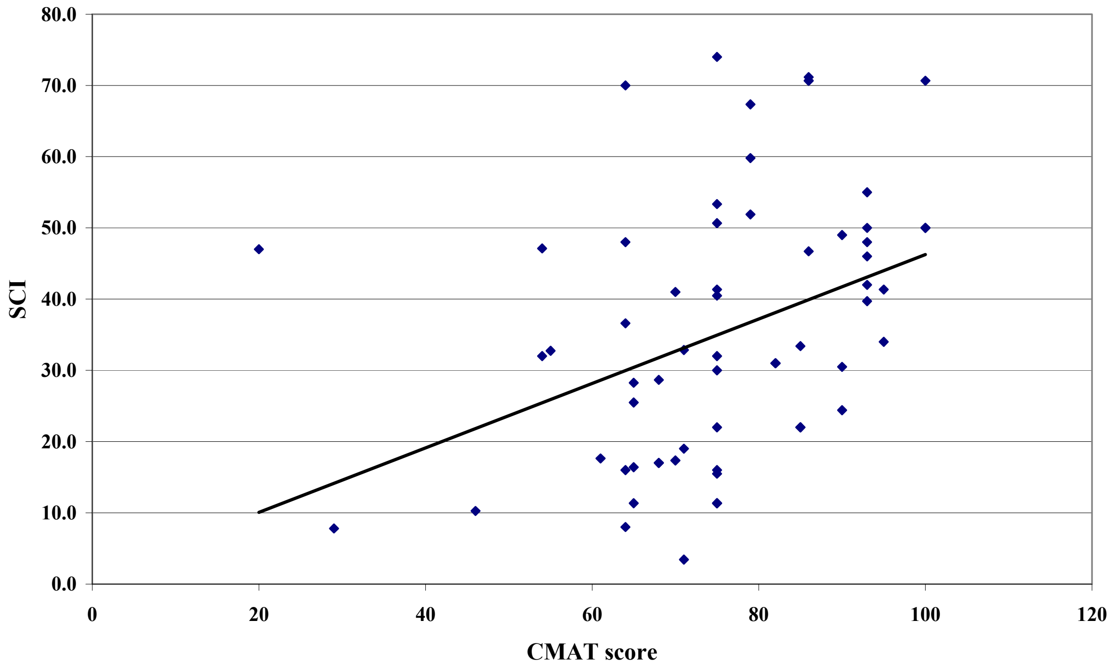


Figure 24 Plot of concept map scores generated using the Structural Complexity Index (SCI) versus CMAT.

Table 16 Regression data for the plot of the Structural Complexity Index versus the CMAT score.

	R^2	Slope	95% Confidence Interval
All Maps	**0.1191**	**0.52 ± 0.18**	**0.17 − 0.87**

The SCI for Student 28's concept map on Concept Mapping Activity 2 is nearly double that of the next closest SCI for any assignment. The reason for the high SCI on this particular concept map is the networked nature the student displayed. While this type of score is indicative of a complex structure, for the purposes of correlating the SCI to the scoring of concept maps in the CMAT program, the data point is omitted. The resulting plot (Figure 25) and regression analysis (Table 16) is shown below.



A plot of the SCI versus CMAT scores after removing the outlying data point from Figure 24. The regression line statistics are shown in Table 17.

Figure 25 Plot of concept map scores generated using the Structural Complexity Index (SCI) versus CMAT with outlier omitted.

Table 17 Regression data for the plot of the Structural Complexity Index versus the CMAT score disregarding the outlying data point.

	R^2	Slope	95% Confidence Interval
All Maps except outlier	**0.1417**	**0.45 ± 0.14**	**0.18 − 0.73**

The data in Figure 25 and corresponding regression statistics in Table 16 indicate that there is little correlation between the CMAT score and Structural Complexity Index for a concept map. Therefore the Null Hypothesis, H_4, is rejected. This result makes intuitive sense, inasmuch that in the CMAT scoring technique the veracity of the propositions was scored, and the SCI does not consider the veracity of the propositions, only how the map is connected together. However, the SCI does take into account the number of propositions created. Therefore, a concept map with few propositions would receive a low score in the CMAT program and a low SCI. In Figure 26 and Table 17, the regression analysis for the separate concept mapping assignments is provided. The results show that concept mapping assignment 1 has a much different correlation between the SCI and CMAT score compared to concept mapping assignments 2 and 3.

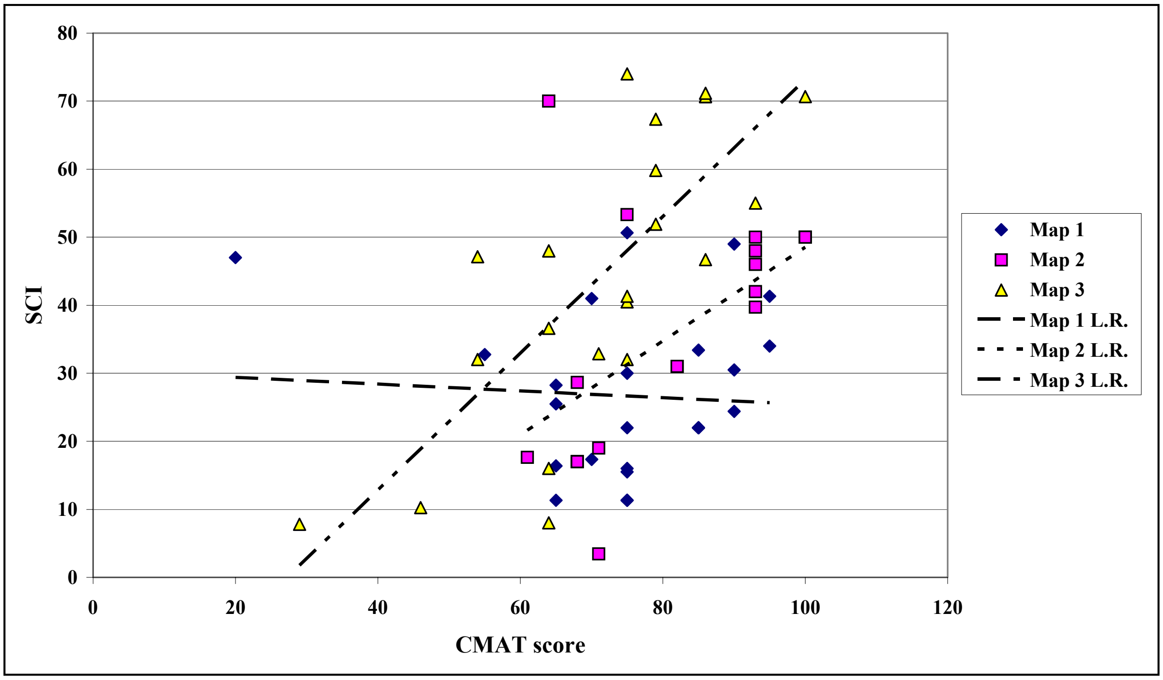


Figure 26 Plot of concept map scores generated using the Structural Complexity Index (SCI) versus CMAT separated by Concept Mapping Assignment..

Table 18 Regression statistics for the comparison of SCI and CMAT scores of concept maps for each individual Concept Mapping Assignment.

	R^2	Slope	95% Confidence Interval
All mapping assignments combined			
All Maps	**0.1417**	**0.45 ± 0.14**	**0.18 − 0.73**
Individual mapping assignments			
Map 1	**0.0041**	**−0.050 ± 0.16**	**−0.38 − 0.28**
Map 2	**0.2989**	**0.69 ± 0.25**	**0.17 − 1.21**
Map 3	**0.5812**	**1.0 ± 0.19**	**0.61 − 1.4**

The data in Table 18 indicates that several factors are at work in the students' performance on the concept mapping assignments:

- Students were not comfortable with concept mapping early in the semester.
- Concept maps with more terms produce a better correlation between the two different scoring methods.

The fact that there is not a significant correlation between a concept map's SCI and CMAT score does not invalidate the SCI as a scoring metric. The SCI is meant to provide an additional score to a concept map to assist the instructor in determining the structural character of the concept map. Since there is not a significant correlation to the relational scoring method employed by CMAT, one conclusion could be that the SCI is independent of the CMAT score. Other observations from experiment 3 help to determine the effectiveness of using the SCI as a concept map scoring method.

Other observations from the data in experiment 3

The most significant factor in determining the usefulness of the SCI in scoring concept maps is how *changes* in concept maps are reflected in the SCI. In Figure 27 and Figure 28, two concept maps drawn by different students are shown. The two concept maps have the same CMAT score, but different SCI scores.

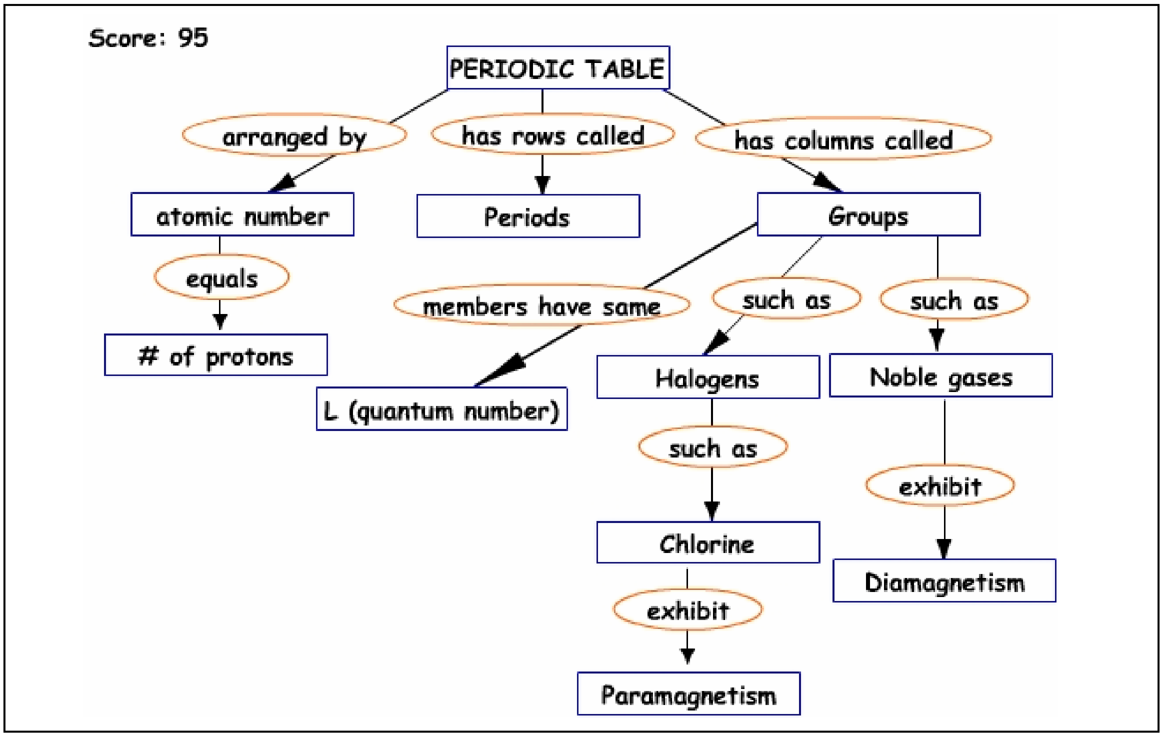


Figure 27 Concept map “A” used for SCI scoring example.

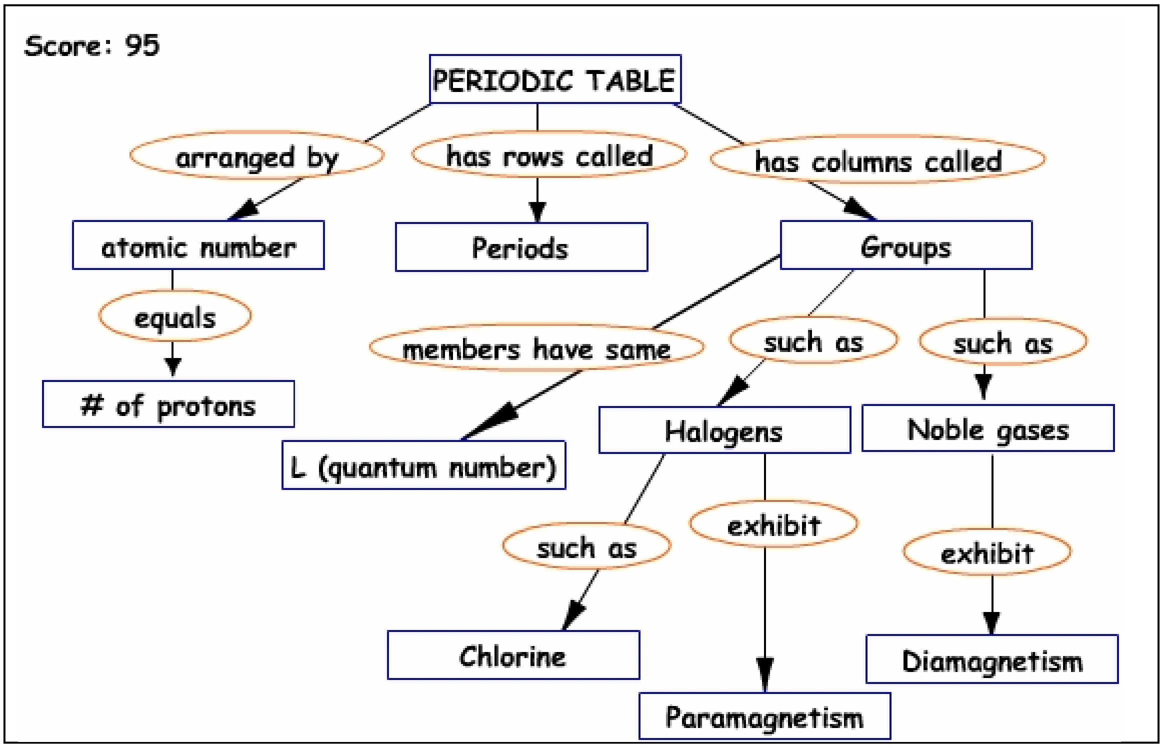
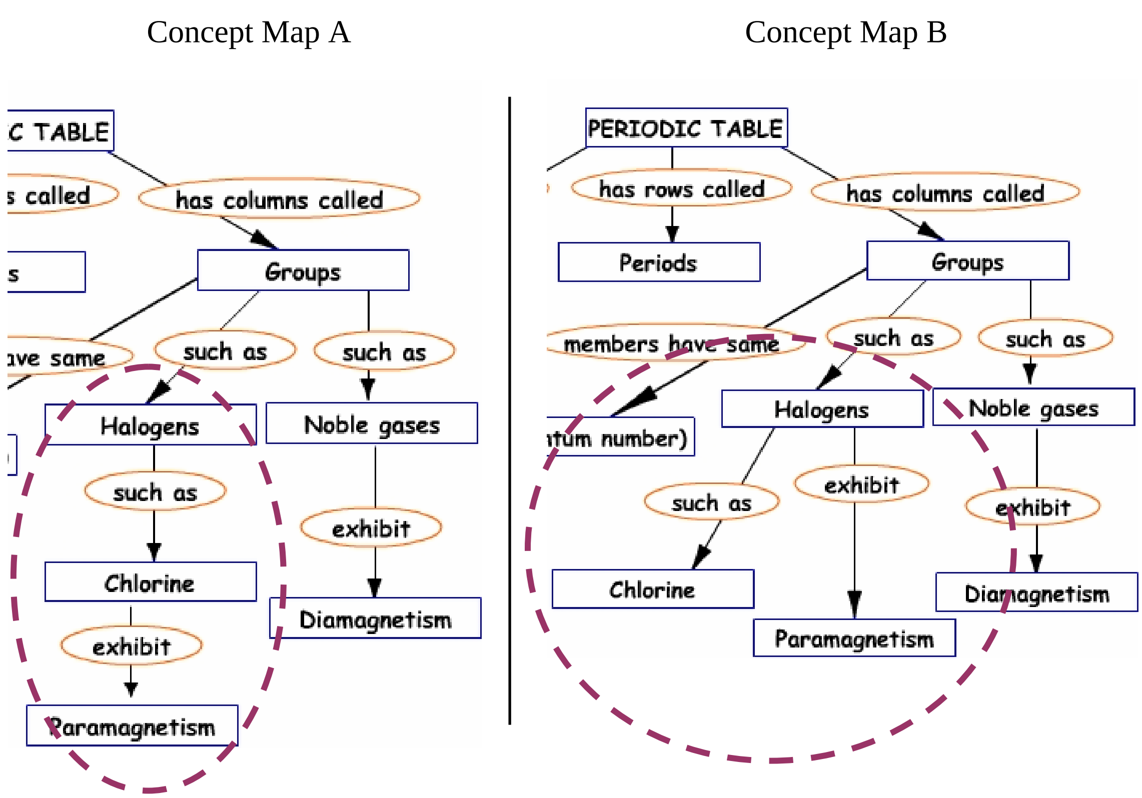


Figure 28 Concept map “B” used for SCI scoring example.

The difference between Concept map A and Concept map B is shown in Figure 29. In Concept map A, the term “Paramagnetism” is connected to the term “Chlorine”, whereas in Concept map B, the term “Paramagnetism” is connected to the term “Halogens” as is “Chlorine.”



Portion of Concept Map A and Concept Map B that is different from one another to demonstration the SCI.

Figure 29 Comparison of structural features between Concept maps “A” and “B.”

The CMAT score for both of these concept maps is the same however, the SCI for Concept map A is 34 and the SCI for Concept map B is 41. The difference results from the extra branching point in Concept map B. In Concept map A, the longer chain resulting from placing “Paramagnetism” after “Chlorine” did increase the average chain length of the concept map, but that alone does not constitute a more structurally complex feature. This example shows the ability of differentiating small variations in concept map structure *via* the SCI.

The next example compares two concept maps with different CMAT scores but similar SCI scores. In Figure 30, Concept map C scored a 100 in CMAT and a 50 for a SCI. In Figure 31, Concept map D scored a 75 in CMAT and a 53 for a SCI.

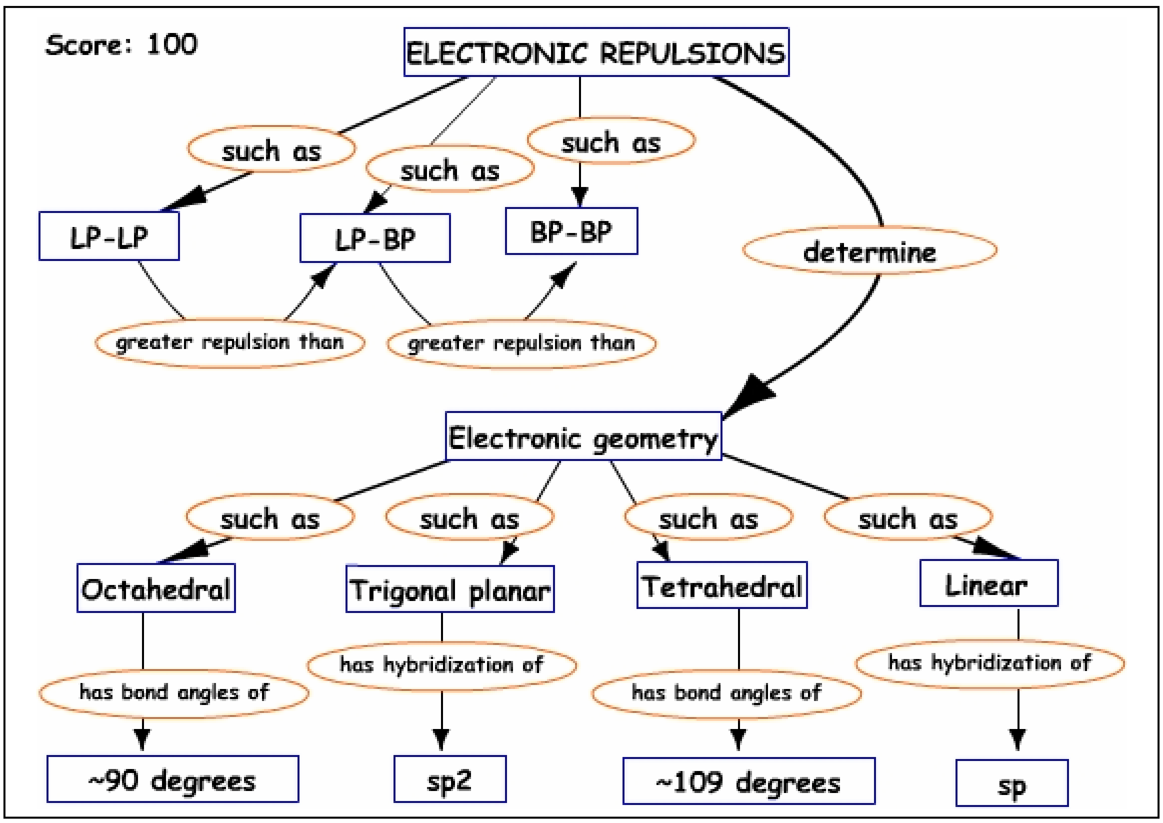


Figure 30 Concept map “C” used for SCI scoring example.

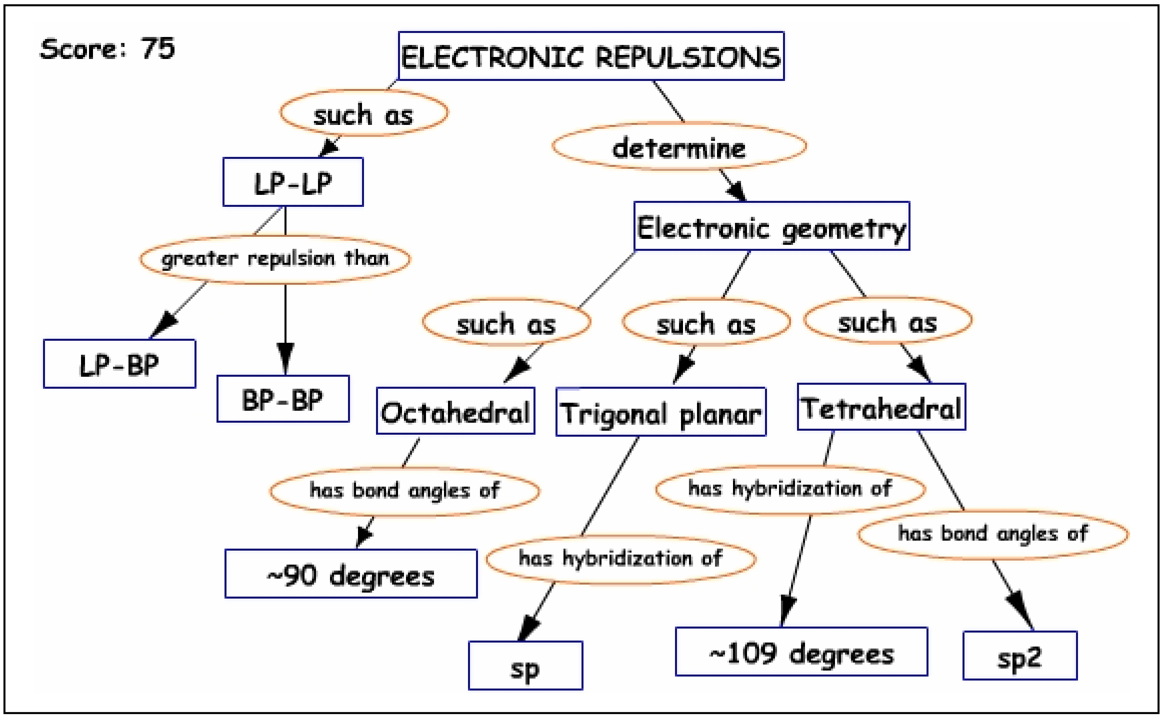


Figure 31 Concept map "D" used for SCI scoring example.

In Concept map D, the student omitted the "Linear" concept and made several more incorrect propositions. These errors caused the CMAT score in Concept map D to be much lower than Concept map C. However, the structure of Concept Map D contains an additional branching point at the "Tetrahedral" term increasing the SCI of Concept map D. The omission of terms – 11 propositions in Concept map D compared to 14 propositions in Concept map C – lowers the SCI for Concept map D, but not to a large degree.

SUMMARY

In summary, the automated scoring in the CMAT program can adequately replicate hand scoring the concept maps. The type of concept mapping assignment used in this study is more confining than an open ended concept mapping assignment; however, if data from a large number of students is available (collected from many open ended concept mapping assignments), the CMAT approach of scoring acceptable propositions is still a viable option. The CMAT approach is similar to the scoring method reported by Francisco and Nakleh in that both approaches score concept maps based upon relational scoring – or the factual correctness of the propositions. The differences in scores generated by the different relational techniques are from the philosophical approach as to whether there is required or necessary number of propositions in the concept mapping assignment.

A different approach to scoring concepts created in this study is based upon the structural characteristics of the concept maps. The SCI demonstrates an ability to differentiate between two concept maps with slight structural differences. In instances where concept maps have different numbers of concepts and some incorrect propositions, differences in SCI are more difficult to interpret. Coupling these observations with earlier results indicating that the SCI is a different metric than the CMAT scoring would suggest that both scores be used to describe a concept map.

Relationship to Bloom's Taxonomy of Cognitive Skills

The relationships between the different concept map scoring schemes and the levels of Bloom's Taxonomy are show in Table 18.

Table 19 Levels of Bloom's Taxonomy included in the different concept map scoring schemes

		Concept map scoring schemes		
		CMAT	Francisco and Nakleh	SCI
Levels of Bloom's Taxonomy	Knowledge	?	?	–
	Understanding	✓	✓	–
	Application	✓	✓	✓
	Analysis	–	–	✓
	Synthesis	–	–	✓
	Evaluation	?	?	?

The knowledge level of Bloom's Taxonomy involves the recognition or recall of facts and information. Since the concept mapping assignments provide students with the concepts and linking words, the recall function involved in the knowledge level is untested by the current study. If students were required to generate their own concepts, then recall could be tested. The focus of this study was the process of forming propositions using terms provided. Since the student could make any number of propositions, the ability of the student to make correct propositions tests cognitive skills in the understanding level of Bloom's Taxonomy. Therefore, both the CMAT and Francisco and Nakleh scoring schemes test for the understanding level. The SCI does not evaluate the correctness of the propositions formed, and therefore does not test the understanding level of Bloom's Taxonomy. The application level of Bloom's Taxonomy requires students to use knowledge in a different application. Since concept mapping is not the normal approach to learning, having students create a concept map requires a transfer of 'book-knowledge' to a graphical representation. The rich nature of concept mapping assignments to include a variety of beginning concepts or examples allows the instructor to assess the student's ability to create useful propositions based upon the new material. The analysis level of Bloom's Taxonomy requires the simplification of complex concepts into simpler parts. Within the relational scoring schemes the ability to break down the more complex concepts could be part of the acceptable propositions, however, that part was used as an example of understanding. The analysis level, as used in this comparison, represents the student's ability to support complex ideas with well structured underlying understanding. Therefore, the SCI would be a measure of the student's analysis ability. A high SCI would indicate the student has the ability to connect several levels underneath a concept, a very high SCI could indicate unnecessary confusion. A lower SCI would indicate a student's lack of ability to connect a complex to other knowledge. The synthesis level of Bloom's Taxonomy requires student's to use existing knowledge to create a product with new meaning or purpose. The two relational scoring schemes test propositions on an individual basis. A connection to the synthesis level could be made for a high score in CMAT or the Francisco and Nakleh approach to show the synthesis level is tested, however, the SCI would seem to contain the necessary test for synthesis. A fragmented concept map may score well in terms of its relational correctness, but the SCI would be very low. As seen in this study, the fragmented concept maps produced the most deviation between automated scoring and handing scoring using the CMAT approach. In the future of automated scoring, this could possibly remain a trouble spot. However, the SCI for these fragmented concept maps showed the maps to be deficient compared to more well formed concept maps. Therefore, to use the concepts in the concept mapping assignments to produce a product with new meaning, the SCI is a better metric than the relational scoring schemes. Finally, the evaluation level in Bloom's Taxonomy requires students to make judgments about the validity of a statement. In terms of a single concept mapping exercise, this ability is not tested. If a student would have the opportunity to revise the concept map, then it would be possible for all scoring methods to test evaluation cognitive skills.

Comparison to course performance

A comparison of concept mapping scores to course performance is not made in this study. The scope of this study is on the automated scoring of concept maps and how concept map scores generated by different techniques correlate to one another. The comparison of grades earned by students to their concept mapping scores is beyond the scope of this study. In addition, a flaw exists in comparing a student's performance on a multiple–choice exam to their concept map score. The concept mapping assignment requires a different cognitive activity than the traditional assessments used in the course where the study took place. Therefore, it is my view that the comparison to course performance, either final grade or individual exam, is inappropriate.

Chapter 5. Conclusions

Overview Of The Study

This study has two phases – a development phase and an experimental phase. During the development phase of the study, two novel tools were created. First, a web based concept mapping program (CMAT) that included an automated scoring component was created. The concept maps created in CMAT were used as the data in the second phase of the study. The second novel tool created during the first phase of the study was the Structural Complexity Index, a concept map scoring method that evaluated structural components of a concept map.

The second phase of the study consisted of three experiments.

- The ability of the CMAT program to score concept maps was investigated by comparing the maps scored in the CMAT program to hand scoring the same maps.
- Scoring concept maps using a different method – reported by Francisco and Nakleh [37] – was compared to the scoring scheme employed by the CMAT program.
- The novel Structural Complexity Index (SCI) was used to score the concept maps created in this study. The relationship of the SCI to the scoring scheme employed by the CMAT program was investigated.

The study used concept maps drawn by the 32 students enrolled in the web based freshman chemistry course (CH 301wb) at The University of Texas at Austin during the spring of 2003. Three different concept mapping assignments were created by the teaching assistant for the course and the completed maps were used by the researcher who did not participate in the teaching of the course. A total of 68 concept maps were used to test the following hypotheses:

- *H1: A concept map's score from CMAT will be the same as the concept map's score from hand grading.*
- *H2: A concept map score generated by CMAT will be the same as the score generated by the Francisco and Nakleh scheme.*
- *H3: A rank ordering of concept maps based upon scores generated in CMAT will be the same as the rank ordering based upon scores generated by the Francisco and Nakleh scheme.*
- *H4: A list of concept maps ordered according to the SCI will be no different than a list of concept maps ordered according the CMAT score.*

The data collected in this study resulted in accepting hypothesis H_1 and H_2, and rejecting hypothesis H_3 and H_4.

INTERPRETATION OF THE STUDY

In an ideal world, teachers would have the tools to assess not only a student's rote knowledge, that is, a mastery of facts and how to do standard calculations in isolation—algorithmic knowledge—but also his/her ability to integrate these simpler elements of knowledge toward an end which the students, or anyone, has not seen. In other words, one aspect of an ideal learning environment is to help students achieve the higher skill levels in Bloom's Taxonomy [11]. Concept mapping has been shown to be an alternative (or additional) assessment tool that addresses the stated concerns but has not been widely employed in classrooms. One goal of this study was to demonstrate how technology – in the form of web based interface – can be incorporated into the concept mapping assessment process. Using a simple concept mapping assignment model – the note card technique [26] – students drew concept maps and had them automatically scored on-line. The first experiment in this study showed that automated scoring was the same as hand scoring concept maps. The second experiment in this study showed that the scoring algorithm employed the CMAT program the same as scoring concept maps with the scoring scheme reported previously [37]. The differences between the two scoring schemes, CMAT and the method reported by Francisco and Nakleh, were the result of basic philosophical differences in necessitating the inclusion of all concepts possible or evaluating students based upon the concepts used in their concept map.

The nature of concept mapping – which includes both a relational, or factual, component and a structural component – leads to the question of incorporating all of the information into an easily accessible score. In experiment 3, the Structural Complexity Index was shown to score a concept map substantively different than the relational scoring scheme employed by the CMAT program. Using this new metric, concept maps can be scored according the structural characteristics. This will now allow an evaluation of a student's growth in their cognitive structure as based upon concept mapping.

RELATING CONCEPT MAP SCORING TO EDUCATION THEORY

In the development of any new assessment technique, a relationship between educational theory and the assessment should exist [14]. The concept mapping assessment methods described in this study are based in the constructivist ideals of Ausubel's Assimilation Theory.

Description of Assimilation Theory

Ausubel's work is based upon his view that the most important determinant of *school learning* is what the student already knows (*vide supra*). A later, refined description of Ausubel's work [41] restates this theme in terms of cognitive structures. Ausubel maintains that meaningful learning occurs when new information is *integrated* into an existing cognitive structure in a non–arbitrary and non–verbatim fashion [41]. The starting point for meaningful learning is the presentation of potential meaningful learning materials that possess logical meaning and can be assimilated into the learner's cognitive structure and which contains some *anchoring ideas* to which the new material can be related. This arrangement of what the successful teacher does in concert with the prepared learner is the operational condition that corresponds with the Ausubel quotation shown earlier. The process of anchoring new learning material to existing ideas present in the student's cognitive structure is the basis of Ausubel's Assimilation Theory. Ausubel recognizes [18] three classes of learning relevant to Assimilation Theory: (1) *subordinate learning*, which can be derivative or correlative; (2) *superordinate learning*; and (3) *combinatorial learning*. Subordinate learning occurs when new information is related to either a more inclusive idea or a conception already present in a cognitive structure. The type of relation and how it affects the existing idea in the student's cognitive structure determines whether this process is described as either *derivative subsumption* or *correlative subsumption*. *Derivative subsumption* describes the case when new ideas are made an extension of, or example of, an idea that exists in the student's cognitive structure; the attributes of the original idea remain unchanged and the new information is incorporated into the cognitive structure beneath the more inclusive idea. *Correlative subsumption* is similar to derivative subsumption in that new ideas are incorporated into the cognitive structure beneath the more inclusive pre-existing idea.

However, in correlative subsumption, the new ideas are an extension, modification, or qualification to the pre-existing idea that makes the meaning of the pre-existing idea more nuanced. Superordinate learning describes the process of a new idea encompassing the meanings of more specific pre-existing ideas. The new, more inclusive idea is incorporated into the cognitive structure at a level above the existing ideas. Combinatorial learning describes the process where a new idea is related to pre-existing ideas but is neither more inclusive nor less inclusive than the pre-existing ideas. In terms of cognitive structure, the new idea resides at the same level as the pre-existing ideas.

Cognitive Structure Variables

As the preceding descriptions of Ausubel's view of meaningful learning indicate, the existing cognitive structure of the learner assumes a predominant role in the integration of new knowledge. Ausubel describes four cognitive structure variables that influence meaningful learning [41]:

- availability
- stability
- clarity
- discriminability

By examining the definitions of these terms, we can get a better idea of the influence cognitive structure has toward meaningful learning.

Availability

The availability of relevant anchoring ideas within the cognitive structure indicates the presence of an anchoring concept that can be related to new knowledge in either a subordinate, superordinate or combinatorial fashion is essential [41]. The relationship of the "availability variable" towards a learner's cognitive structure requires that certain concepts and ideas must be present to either (1) proceed with meaningful learning or (2) provide evidence that meaningful learning has occurred. These “available” anchoring concepts will have a unique set of attributes defined by the subject matter or the particular topic within a domain. It is important to note how the location of these anchoring concepts within the hierarchy of the cognitive structure relates to meaningful learning. General, more inclusive anchoring ideas must be available in the learner’s extant cognitive structure to facilitate the meaningful learning and retention of more specific ideas. Therefore, when evaluating a cognitive structure—a concept map—the presence and location of key anchoring ideas must be considered.

Stability

The stability of a cognitive structure refers to the availability of concepts in the cognitive structure over time. More stable ideas and concepts in the cognitive structure are retained longer [41]. Stability does not necessarily relate to the correctness of the idea or concept. Indeed, misconceptions learned very early can become part of a stable cognitive structure and do not become unlearned or forgotten over time [26]. When evaluating the stability within a cognitive structure, specific concepts must be considered. It would be of little value to attempt to assess the stability of the whole cognitive structure because of its ever-changing nature. However, determining when a specific concept enters the cognitive structure, when that concept is related to other concepts, and whether that concept remains active and available in the cognitive structure give a much better determinant of the stability variable.

Clarity

The clarity of ideas and concepts present in the cognitive structure is defined by the degree of explicitness, lucidity, and freedom from vagueness and ambiguity [41].

Clarity as a cognitive structure variable relates the mastery the individual possesses of any particular idea. A learner can demonstrate clarity in a cognitive structure in concept mapping by associating many correct propositions with a specific concept. However, the value of making many trivial propositions (e.g., examples) must be balanced with making a few meaningful propositions to other more general, or inclusive concepts, which are further used in propositions.

Discriminability

The discriminability between two closely related concepts is typically influenced by other cognitive structure variables such as clarity and stability [41]. Only when anchoring ideas are stable and clear is similar knowledge able to be integrated into the existing cognitive structure with its own identity. When similar new knowledge is integrated into the cognitive structure that is unstable or where the anchoring ideas are unclear, ambiguous or competing meanings tend to form [41]. The lack of discriminability between these two similar concepts leads to a lower retention over time of the newly learned material. A reflection of the discriminability between two concepts or a group of related concepts in a concept map is the correct usage of directional links between concepts. Discriminability can also be demonstrated by the correct usage of similar concepts in a single proposition, which can result in an interconnected structure and has been used by several authors as an important indicator of meaningful learning [17, 26, 27].

If the cognitive structure is clear, stable, and suitably organized, precise and unambiguous meanings *emerge* and tend to *retain* their dissociability strength or availability [41]. If on the other hand, the cognitive structure is unstable, ambiguous, disorganized, or chaotically organized meaningful learning and retention is inhibited.

The degree to which these variables can be measured reflects the ability to measure meaningful learning.

Relationship to Concept Mapping

Ausubel's Assimilation Theory describes a *process* of learning. Looking at the cognitive structure variables identified by Ausubel, we see that all of them, in some way, can be thought of in terms of functions over time. As a result, within the cognitive structure, meanings tend to *emerge* and, when learned meaningfully, *retain* their dissociability strength. The key to the assessment of meaningful learning is encoded by these two ideas—emergence and retention.

An assessment of meaningful learning as described here lies in the identification of the *emergence* of key terms, the *placement* of these terms in the knowledge structure and the *progressive differentiation* about those key terms as learning continues.

The emergence of key terms reflects the "availability" variable (*vide supra*). As the key term is further defined (clarity) and assumes a more stable position in cognitive structure, the future linking of related ideas (discriminability) is possible. As successive maps are drawn, the subsumptive, superordinate, and combinatorial learning processes can be identified.

SCORING CONCEPT MAPS ACCORDING TO ASSIMILATION THEORY

Novak and Gowin's original scoring rubric [26] can be related to the four cognitive structure variables of availability, stability, clarity and discriminability. The hierarchy component can be related to the *availability* of a concept. The availability of a concept is determined by its location within the cognitive structure. More general, inclusive concepts should be higher in the cognitive structure so they are available for subsumptive learning. The hierarchy component of the target scoring rubric allows a measure of the levels in the cognitive structure and hence, an indication that some of the higher level concepts are more available. However, a judgment must be made of whether concepts residing in the upper levels of the concept map represent the necessary more inclusive concepts in the subject domain. The proposition component in the scoring rubric can be related to the *clarity* of a concept. Since the hierarchy component of the scoring rubric is influenced by the correctness of the propositions comprising the different hierarchies [40], the hierarchy component is, in turn, also related to the *clarity* of a concept. The cross-link component of the scoring rubric is a proposition and, therefore, related to the *clarity* of a concept. Since a cross-link is formed between concepts at the same hierarchal level, a cross-link could also be related to the *discriminability* between concepts. Novak and Gowin claim that cross-links show a reconciliatory process that allows different concepts to be brought together. The example component of the scoring rubric is again related to the *clarity* of a concept because the examples help define the concept within the learner's mind. Within Novak and Gowin's scoring rubric, the clarity of cognitive structure is greatly emphasized, with availability and possibly discriminability to a lesser extent. Therefore, the original scoring scheme provided by Novak and Gowin may not represent meaningful learning.

The cognitive structure variable *stability* is neglected by the Novak and Gowin scoring rubric. It is evident that a single concept map cannot provide information necessary to determine if meaningful learning is taking place. Therefore, we must shift our focus to analyzing many concept maps drawn over time. This process requires a fundamental shift in the current approach to teaching and assessment, but if we are to measure meaningful learning as Ausubel has defined it, we must give students time and the motivation to revisit their conceptual framework [44].

USING CMAT TO ASSESS MEANINGFUL LEARNING

The scoring of concept maps described in this study – the relational scoring scheme employed by CMAT and the Structural Complexity Index – are representative of this understanding of Ausubel's Assimilation Theory as it applies to concept mapping. The veracity of propositions as scored by the relational scoring scheme employed by CMAT can be used to describe the *clarity* of cognitive structure. The sensitivity of the Structural Complexity Index shown in this study to the addition of concepts in cognitive structure measures the *discriminability* shown by the student. Also, the Structural Complexity Index factors cross-links into its formula in terms of the average chain length. Future studies requiring students to produce and maintain a concept map throughout the learning process. In doing so, these studies will be able to address the questions on how concept map scoring changes with respect to changes in cognitive structure. By tying the relational score and structural score to particular concepts in successive concept maps drawn by a student, the assessment of meaningful learning can be achieved.

FINAL SUMMARY

This study showed that a web based concept mapping program can effectively score student drawn concept maps. The relational scoring scheme employed by the web based concept mapping program scores concept maps similarly to a previously reported method. The newly created Structural Complexity Index evaluates concept maps in a different manner than the relational scoring schemes and, if used in conjunction with the relational scoring method, can be used to assess meaningful learning.

Appendices

APPENDIX A. CONCEPT MAPPING INTRODUCTION

How to Draw a Concept Map

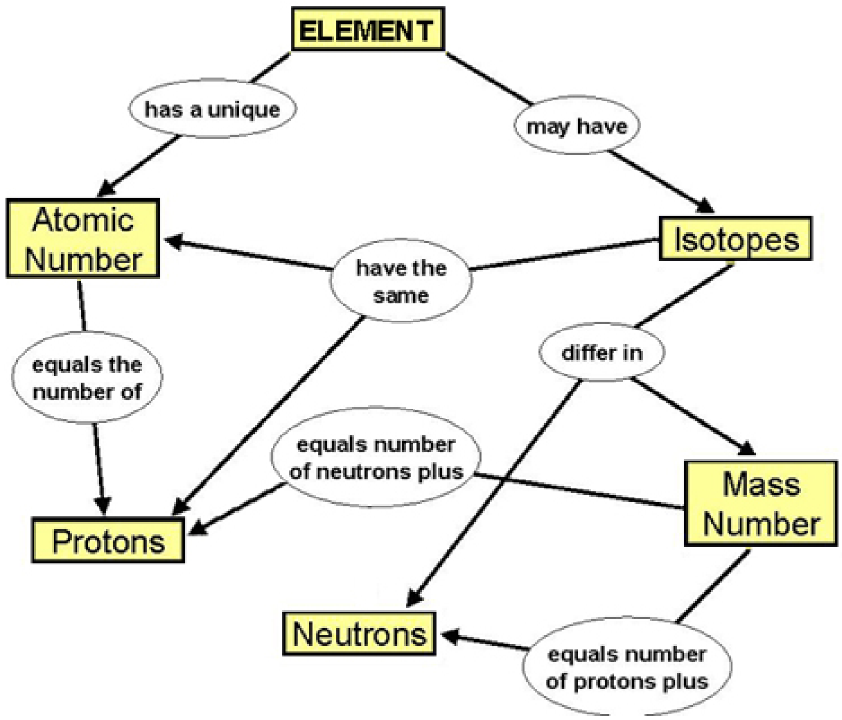
Components of a concept map: Concepts, Linking Words, and Arrows

You will be given a list of concepts, and a list of linking words. The main concept will be capitalized. The concepts will be connected through the linking words with arrows. The connections are directional.

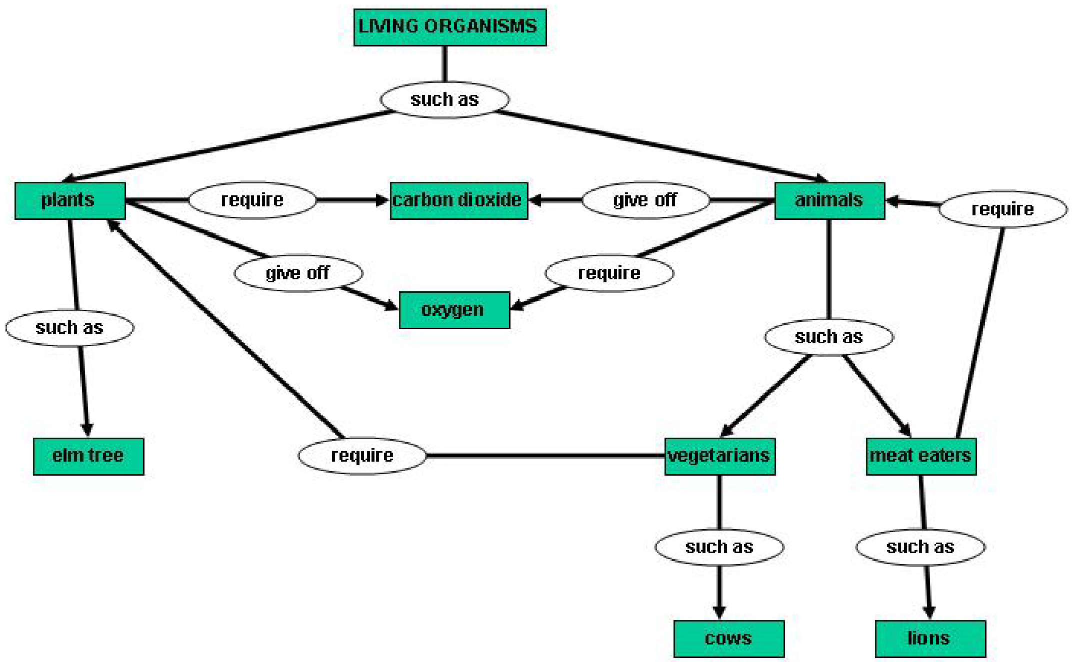
Rules for Construction:

- Each concept from the concept list may only be used once.
- Linking words may be used more than once
- Connect the concepts with a directional arrow and a linking word to make a meaningful relationship
- Concepts are connected from most general to more specific (usually)
- Concepts may have multiple arrows coming from them
- Concepts may have multiple arrows pointing to them
- The capitalized word should be the root word at the start of the "tree".
- Examples: (there are 2 so make sure you scroll down).

Example Concept Map 1



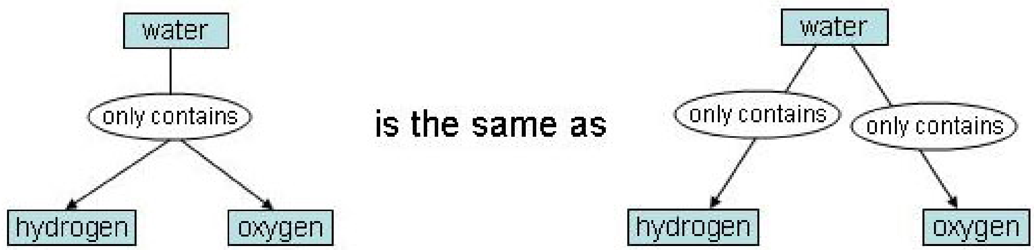
Example Concept Map 2



How Concept Maps are Graded

Grading a concept map is a very subjective task. So it's important that you understand how to "play the game" or in other words, know good mapping strategies.

You can only make binary relationships. For example:



BOTH of these examples make the incorrect statements that (water only contains hydrogen) and (water only contains oxygen). They do NOT say that (water only contains hydrogen and oxygen).

Use the physical meaning of 'consist of' or 'contains'. Many of you put (chemical equations consist of compounds) when it would be more correct to say that (chemical equations consist of symbols) (symbols represent compounds). If I were to physically put two compounds together, say salt and water, I would have salt water, not a chemical equation. But if I were to physically put the appropriate symbols together, I would have a chemical equation.

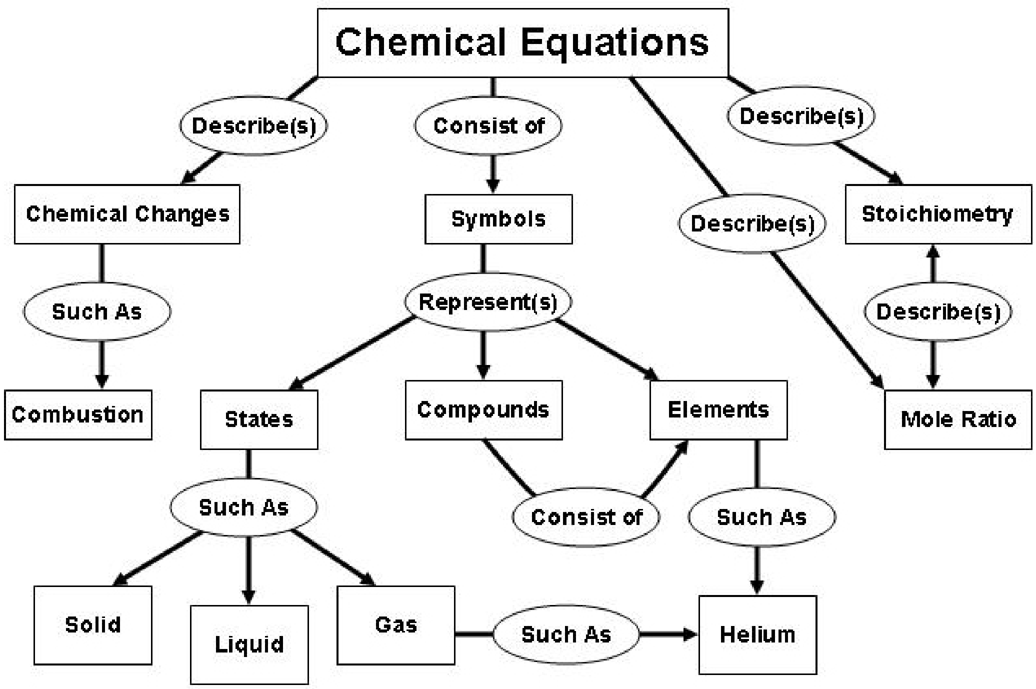
Always use 'such as' for examples. Consider the concept map shown below, if a student used (states consist of gas) (states consist of liquid) (states consist of solid) I know what the student means, but the correct way of saying this is (states such as gas).

Be very particular about the linking words you use!!

Do not make redundant or incorrect statements.

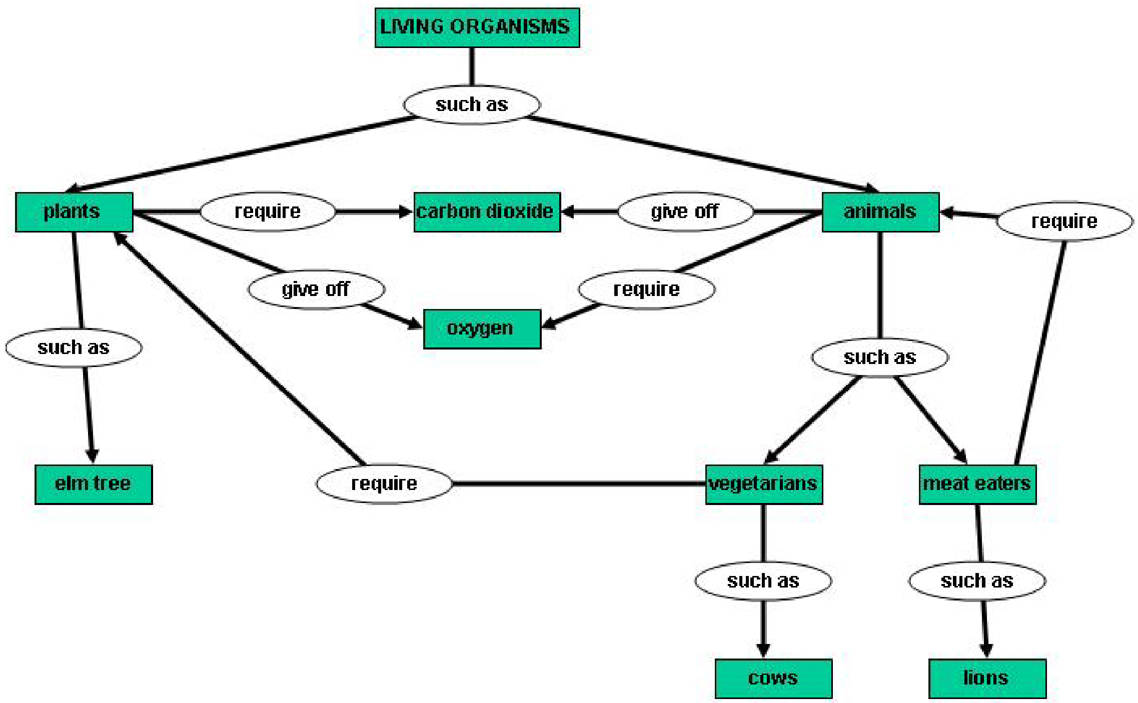
Take a look at the map I made from the exam1 concepts and linking words. It may help you in constructing your next map.

Exam 1 Concept Map



Redundant and Incorrect Statements

Do not make redundant or incorrect statements. Look at the map below. It does not contain any redundant or incorrect statements.



An example of a redundant statement would be *'elm trees require carbon dioxide.'* Even though this is correct, it is redundant because we've already indicated that plants require carbon dioxide and we've already designated an elm tree as a plant.

It is also redundant to say things like *'living organisms such as meat eaters'* or *'living organisms such as elm trees'* or *'animals such as cows'*.

It is correct to say that *'vegetarians require plants'*. It would be incorrect to say that 'animals require plants' since that is only true for certain types of animals.

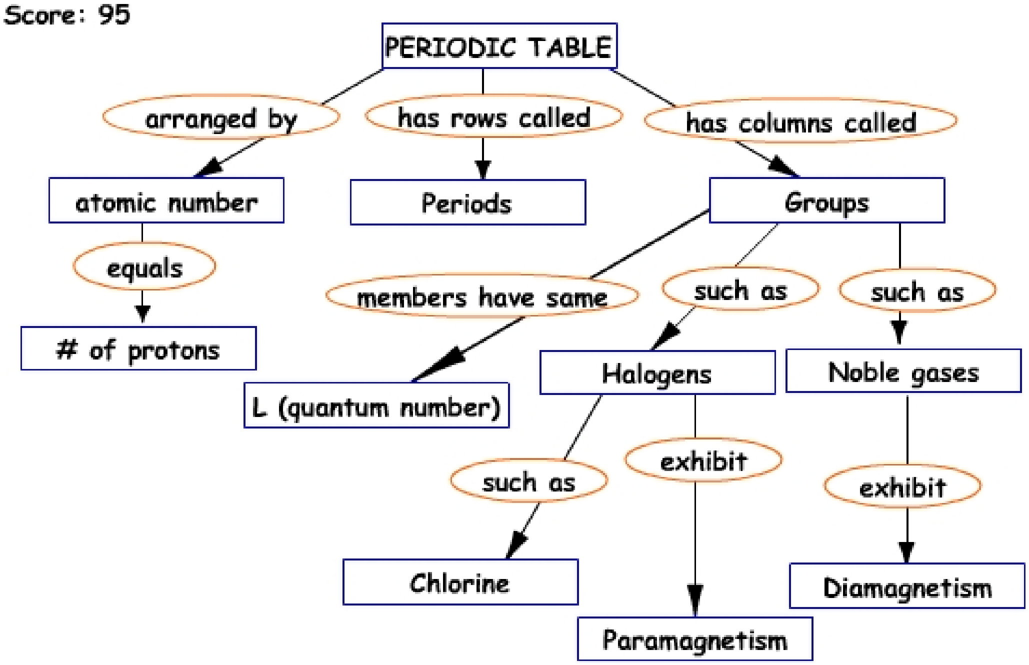
Here's a complicated one... If you added the link *'cows require plants'* to the current map, it would be redundant. BUT, if you removed the statement that *'vegetarians* *require plants'*, the first statement is no longer redundant. HOWEVER, if you did this, you would be missing the idea that ALL vegetarians require plants and not just cows.

APPENDIX B. CONCEPT MAPPING ASSIGNMENT 1

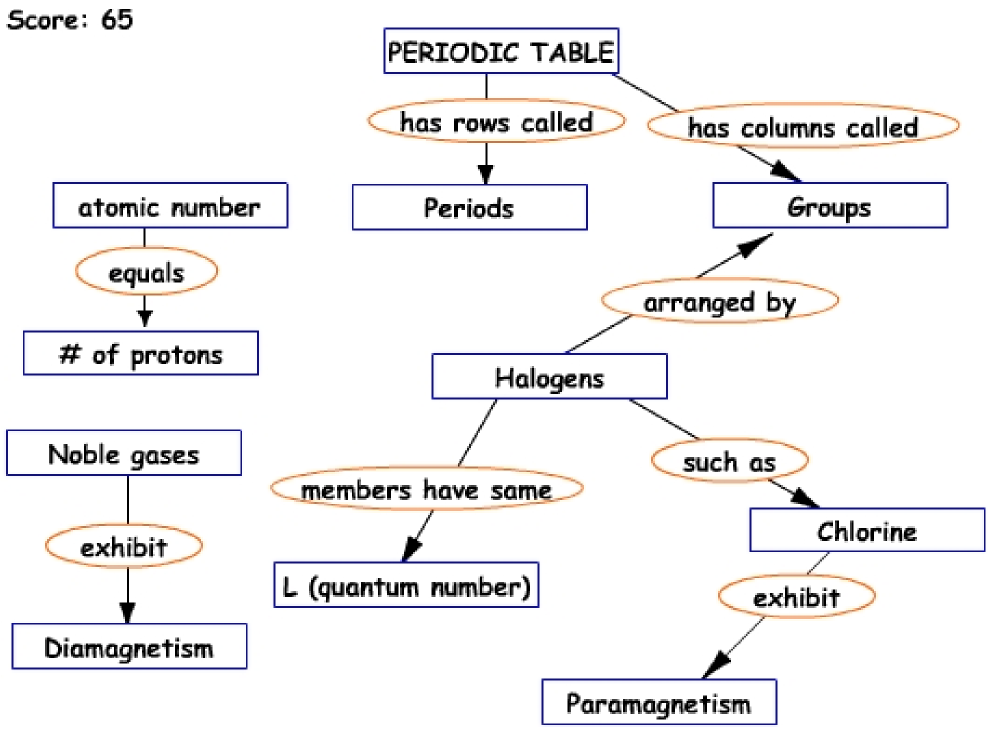
Concept maps submitted by students for Concept Mapping Assignment 1 were drawn in CMAT and stored as a list of propositions in a text file. The submissions for the 32 students enrolled in the CH 301wb course were redrawn into concept maps used for the hand grading procedures in this study. Students who did not submit a concept maps are noted as "*No concept map submitted*."

Student 1

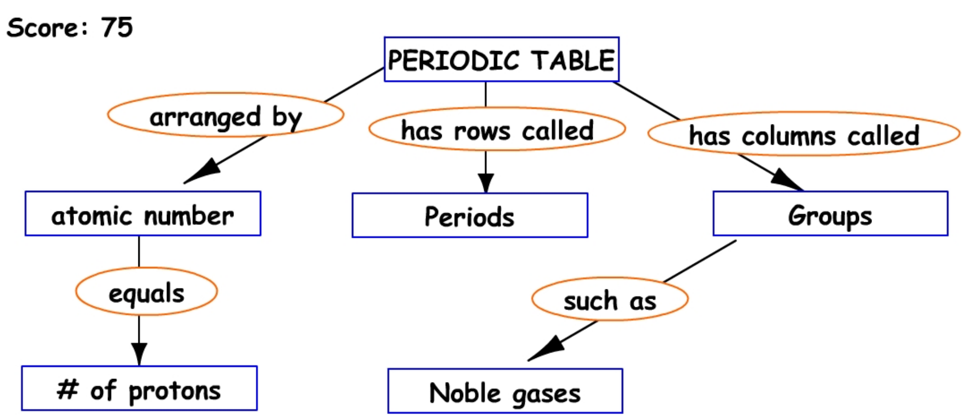
Score: 95



Student 2

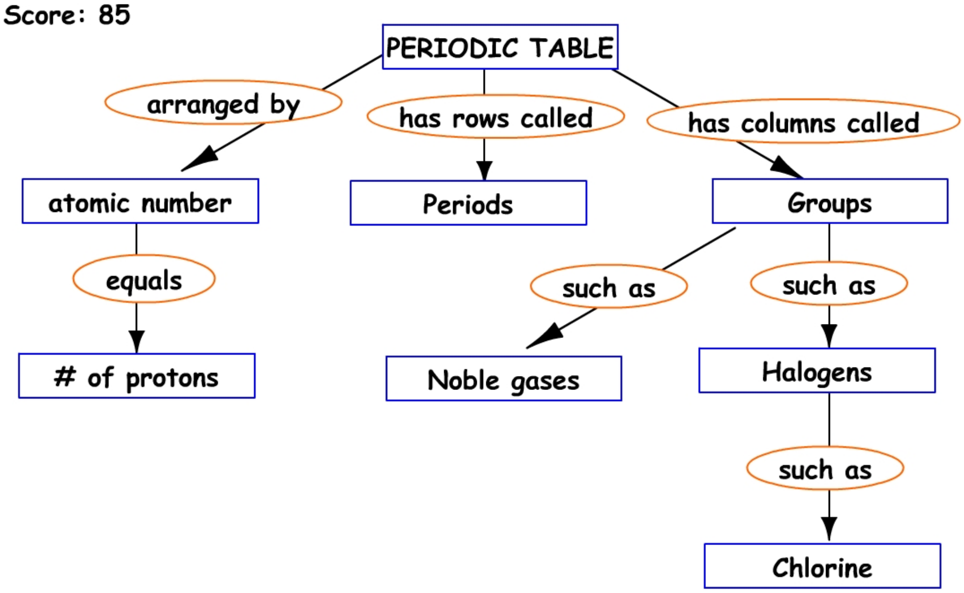


Student 3



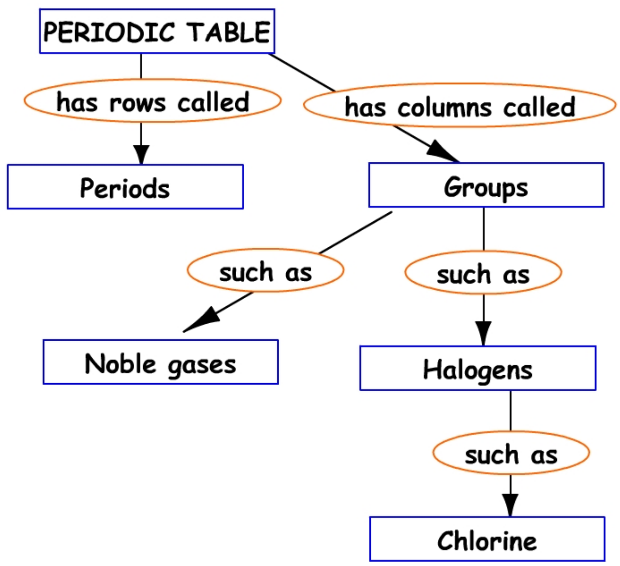
Student 4

Score: 85

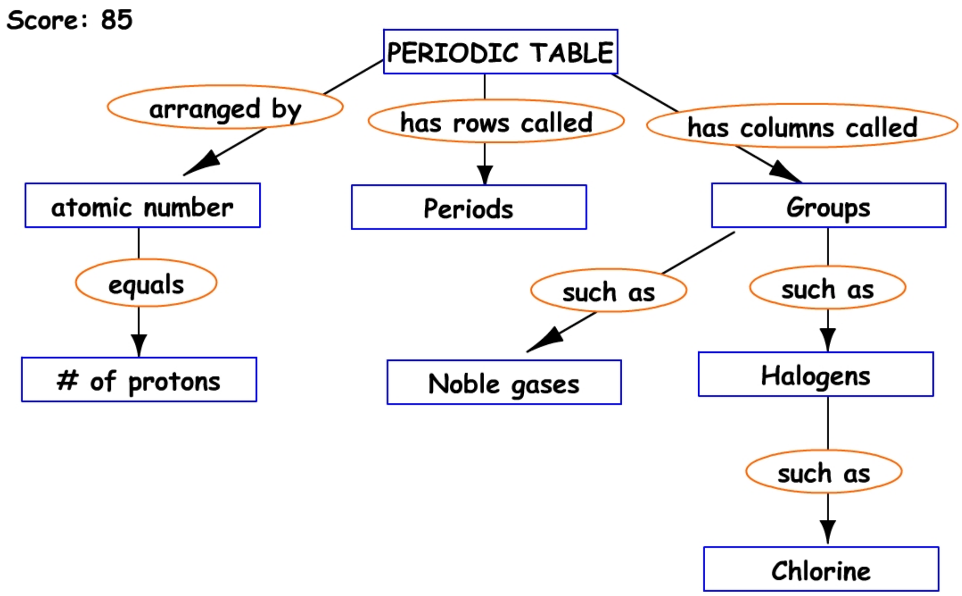


Student 5

Score: 75



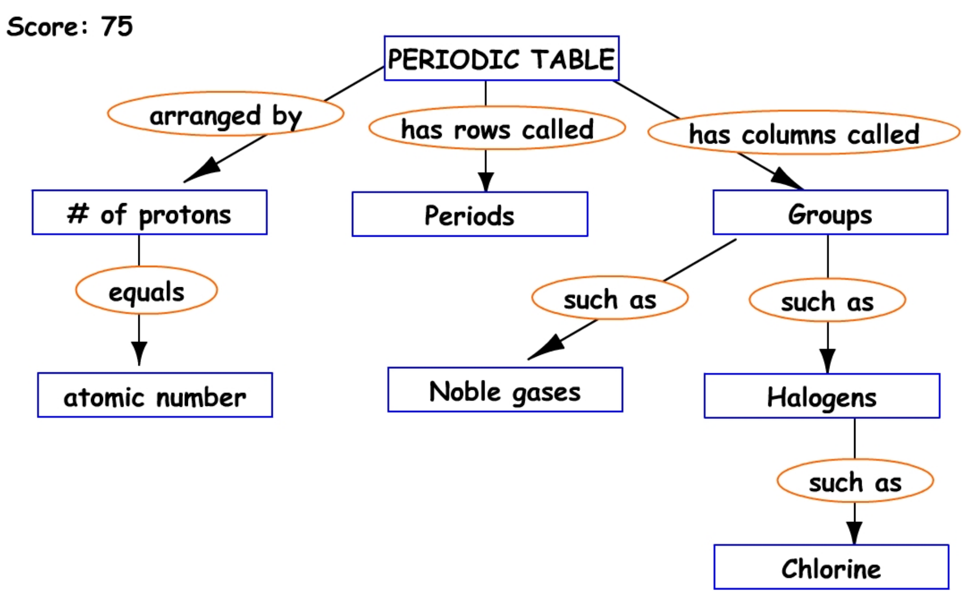
Student 6



Student 7

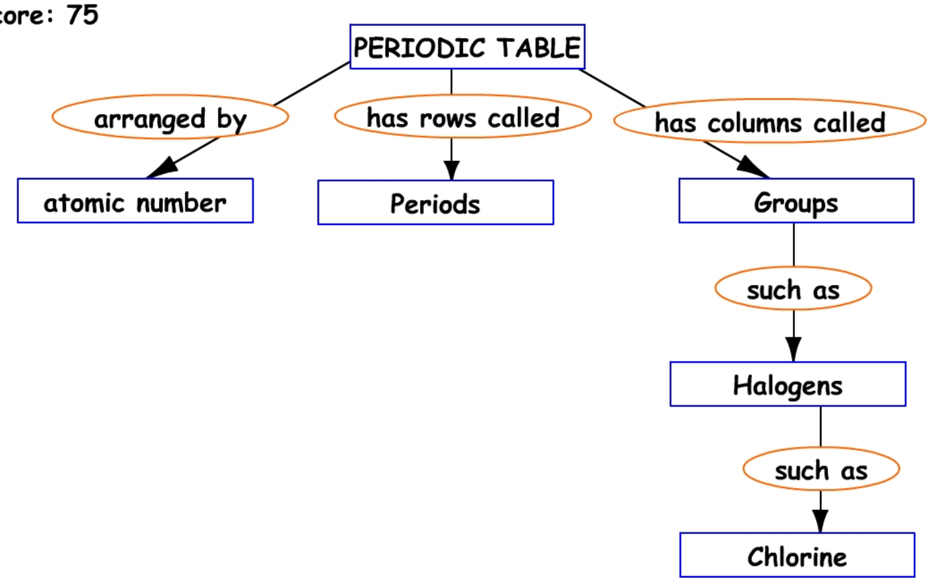
No concept map submitted

Student 8



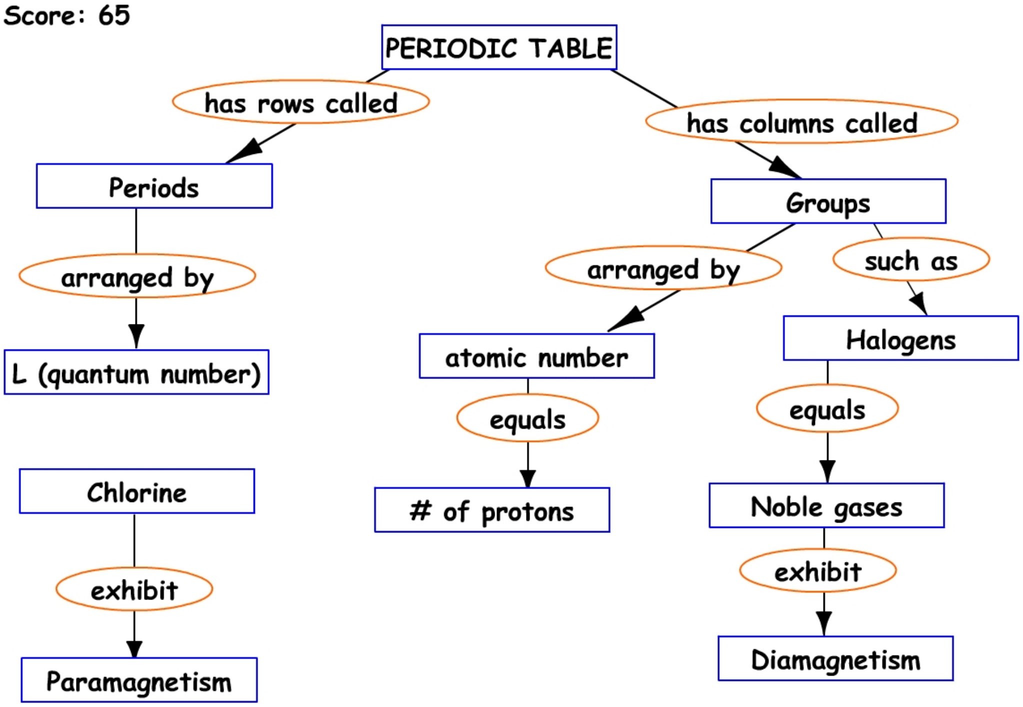
Student 9

Score: 75



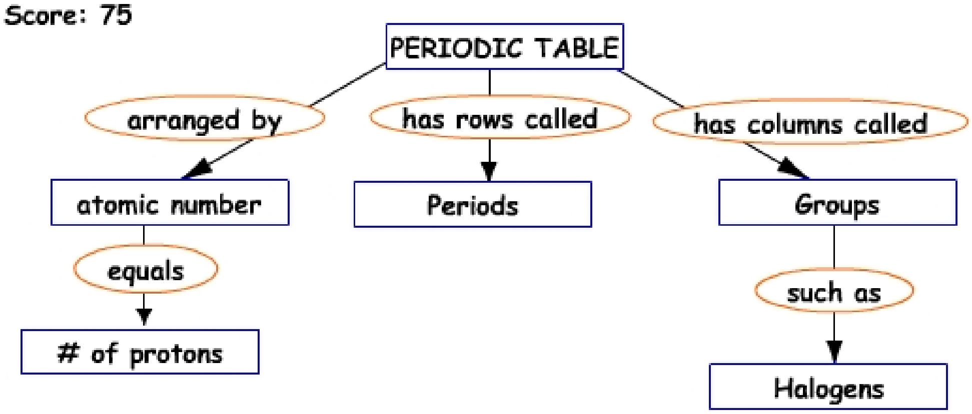
Student 10

Score: 65



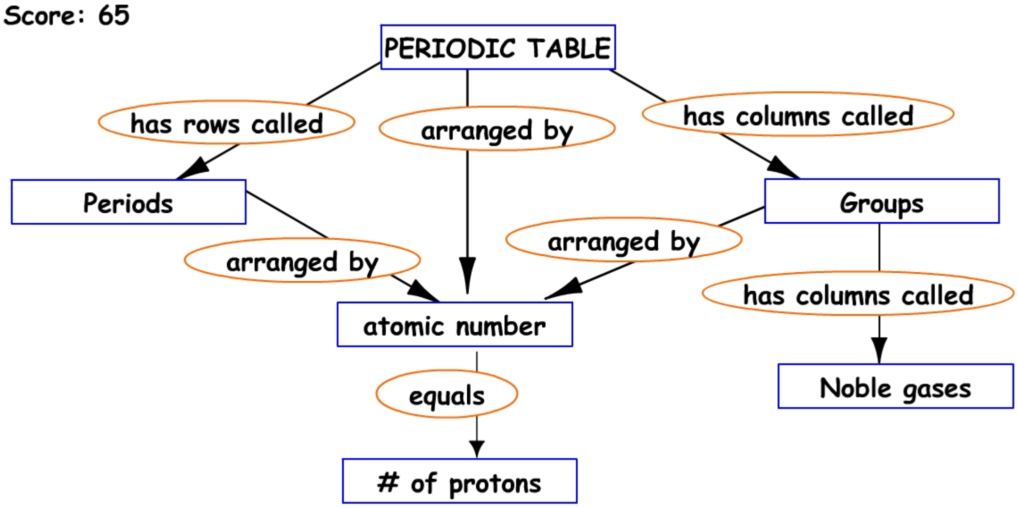
Student 11

Score: 75



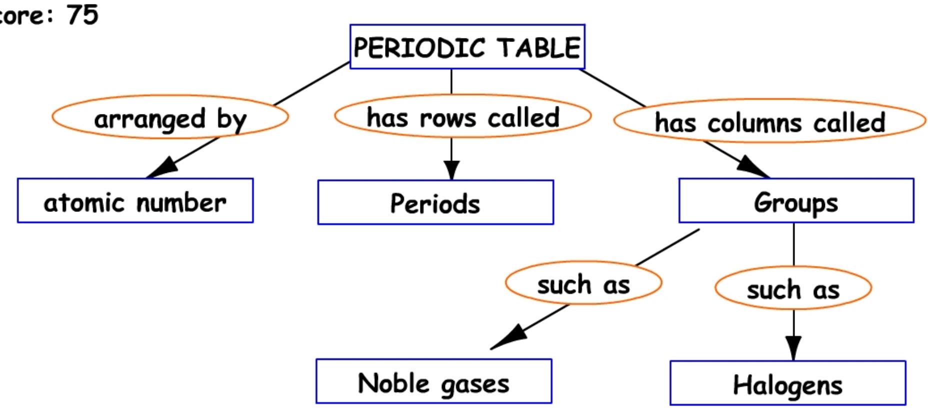
Student 12

Score: 65



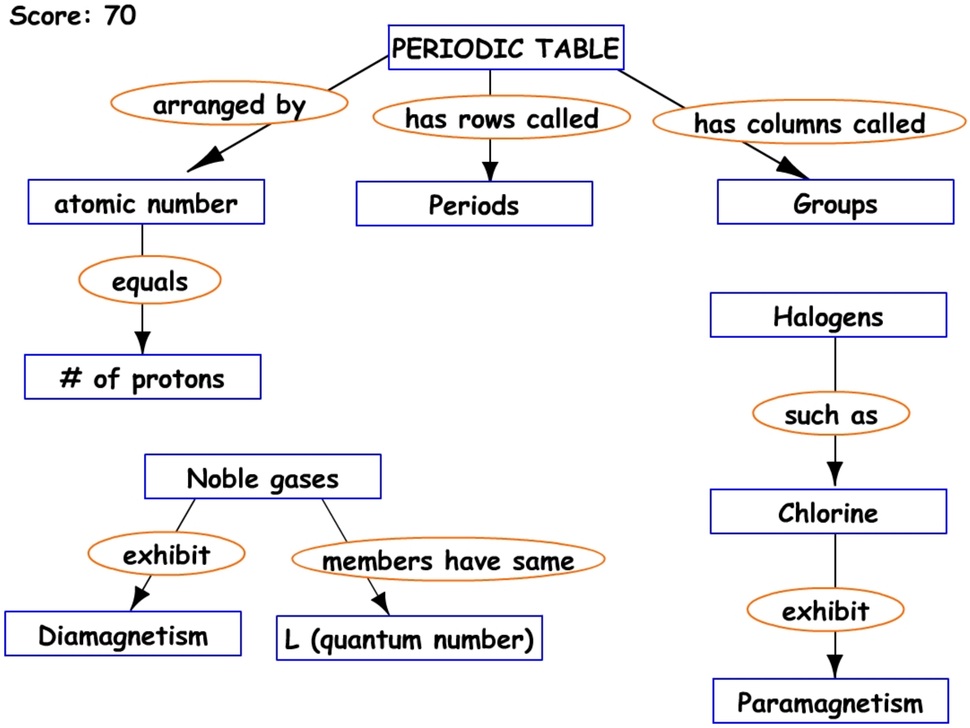
Student 13

Score: 75



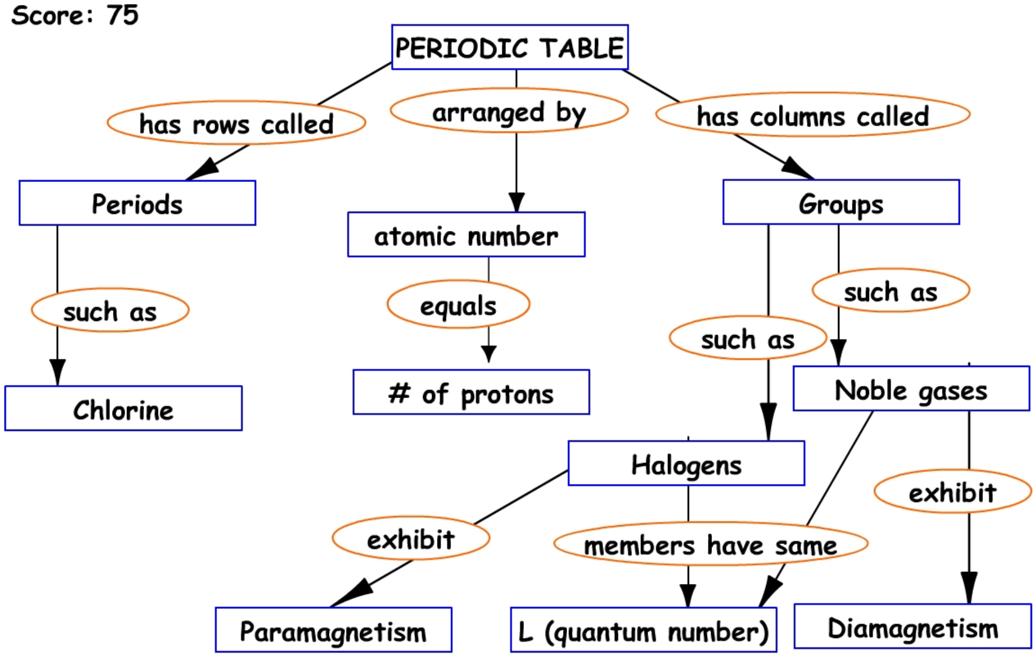
Student 14

Score: 70



Student 15

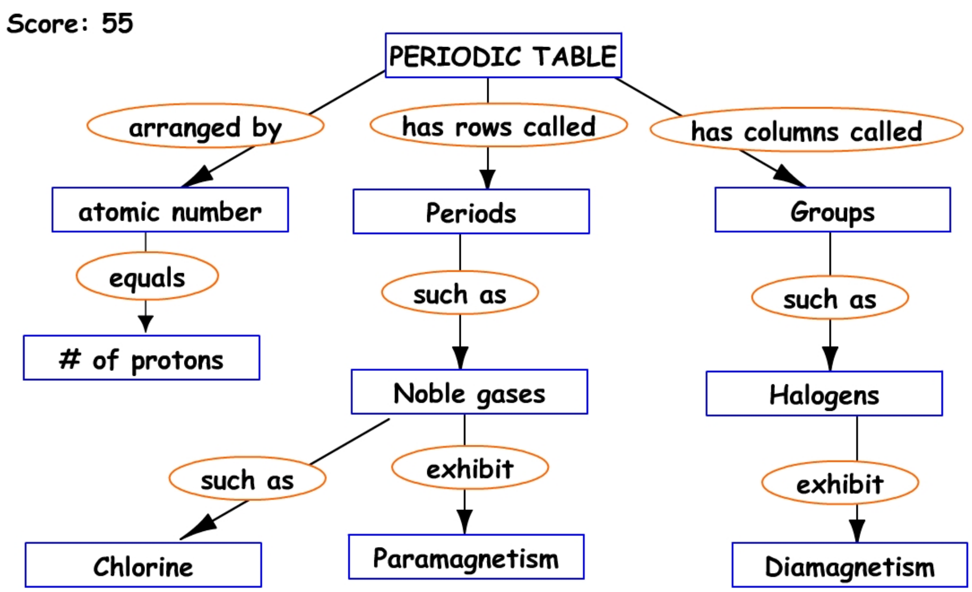
Score: 75



Student 16

No concept map submitted

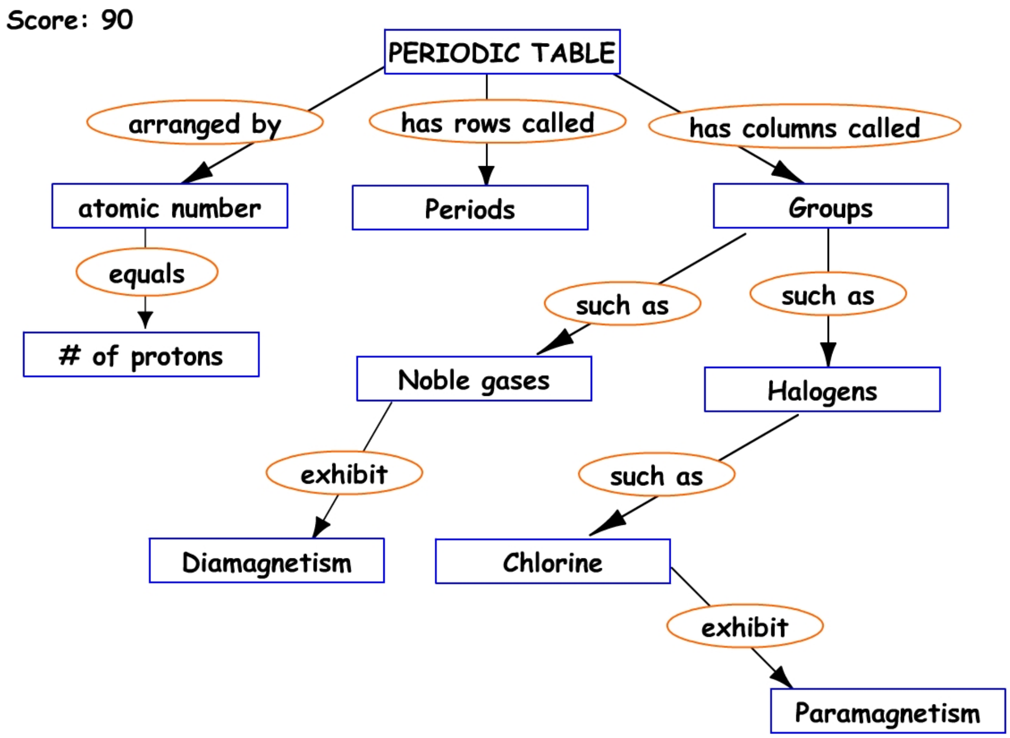
Student 17



Student 18

No concept map submitted

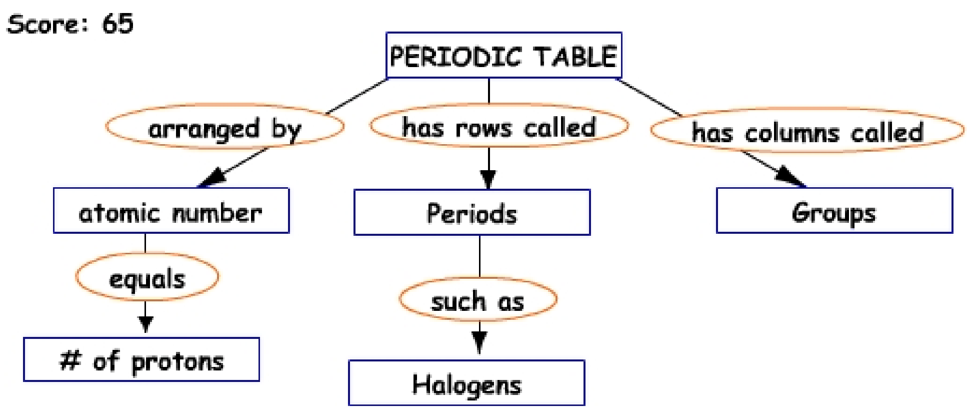
Student 19



Student 20

No concept map submitted

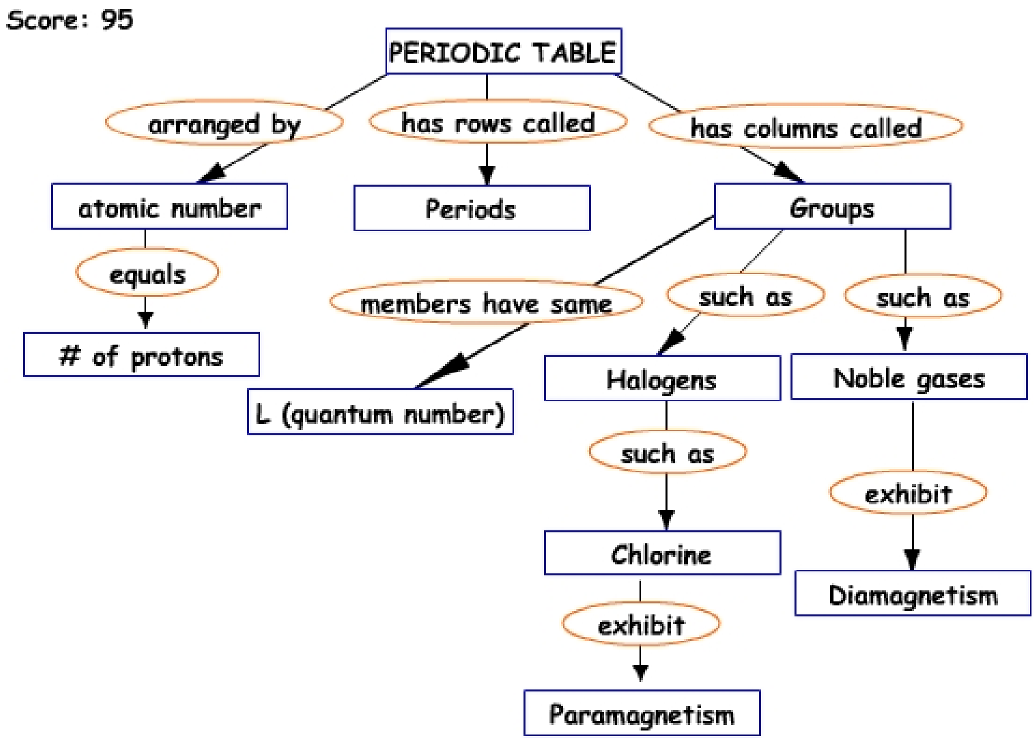
Student 21



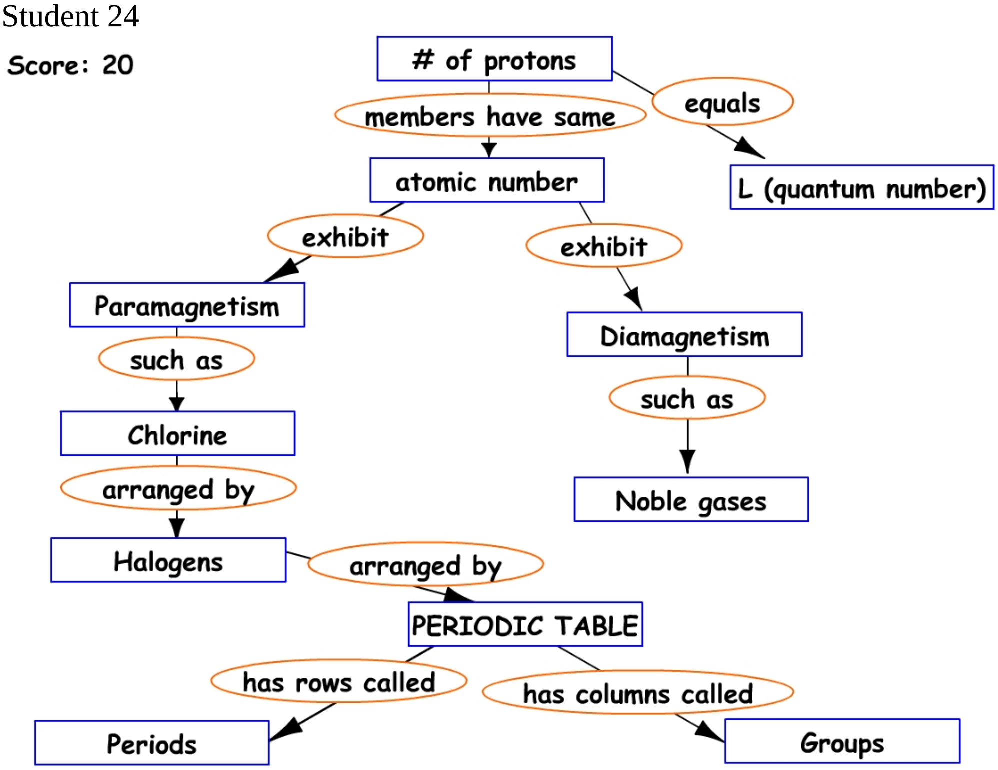
Student 22

No concept map submitted

Student 23

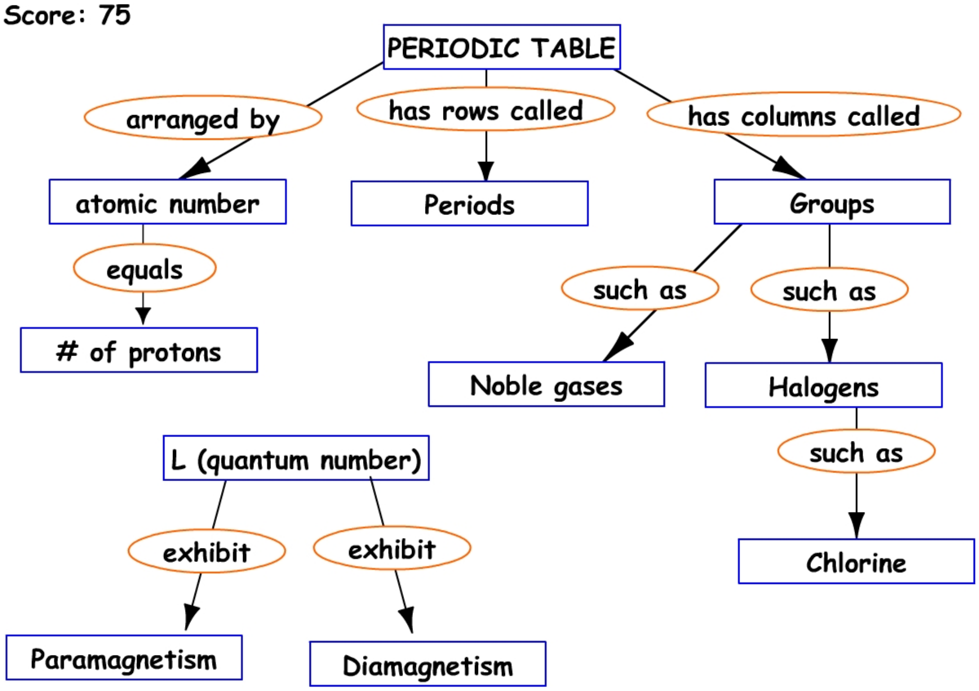


Student 24



Student 25

Score: 75

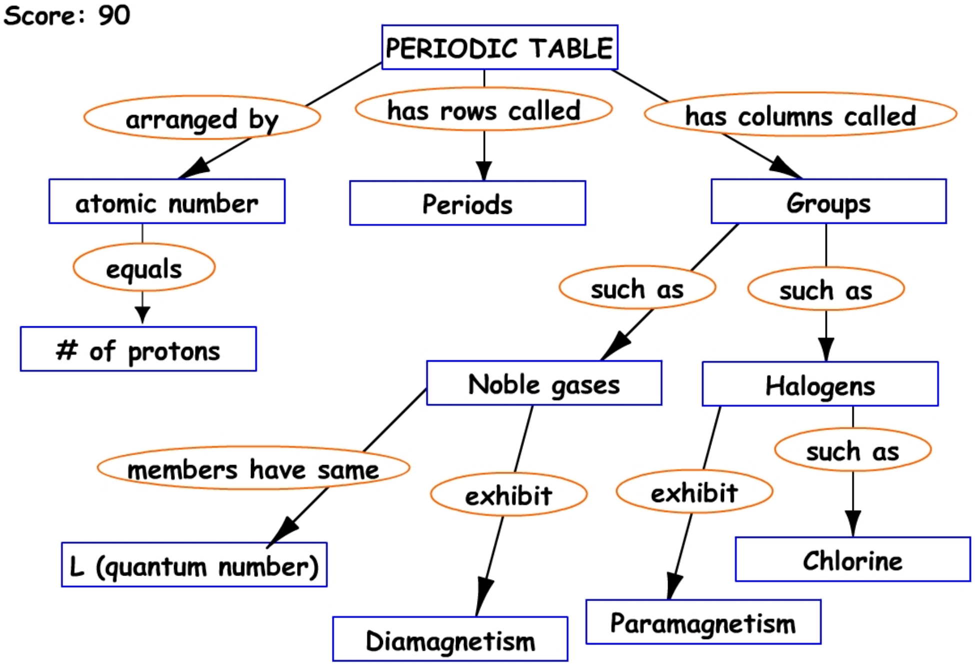


Student 26

No concept map submitted

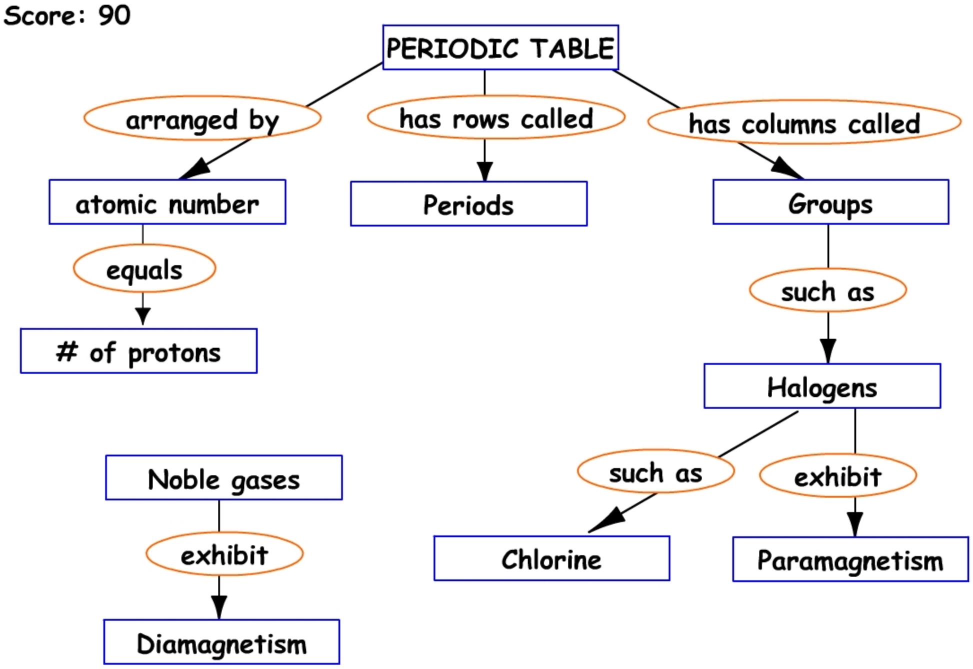
Student 27

Score: 90



Student 28

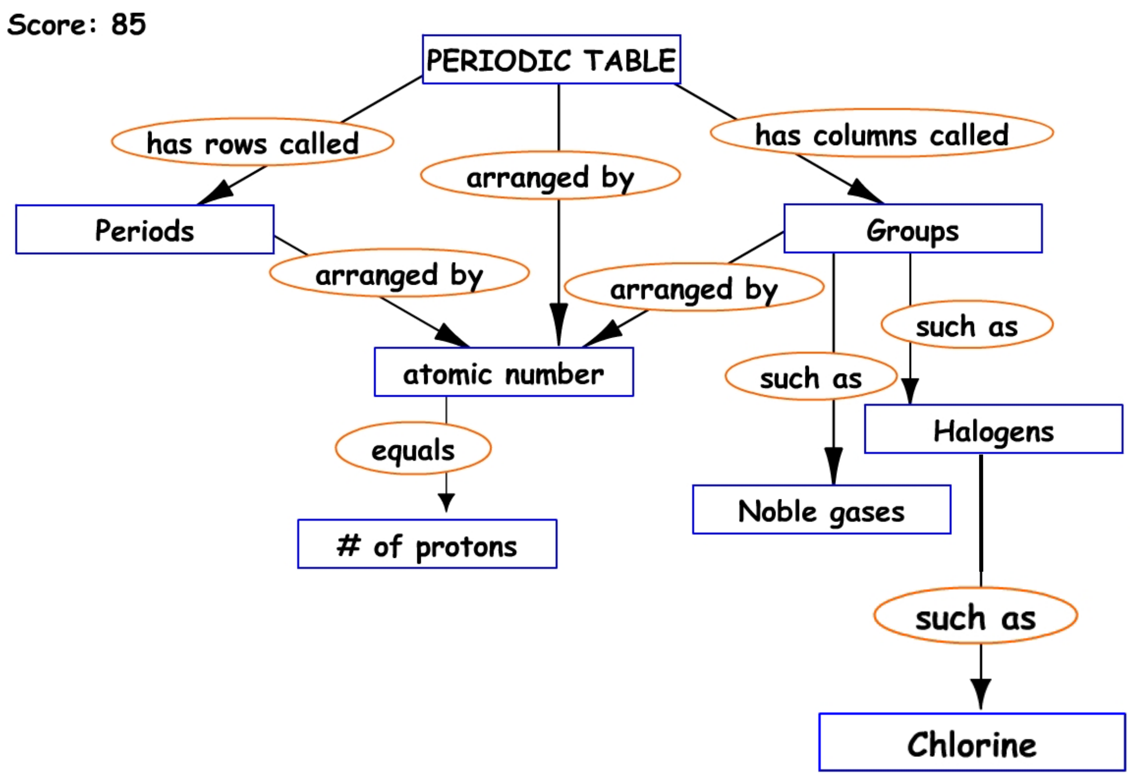
Score: 90



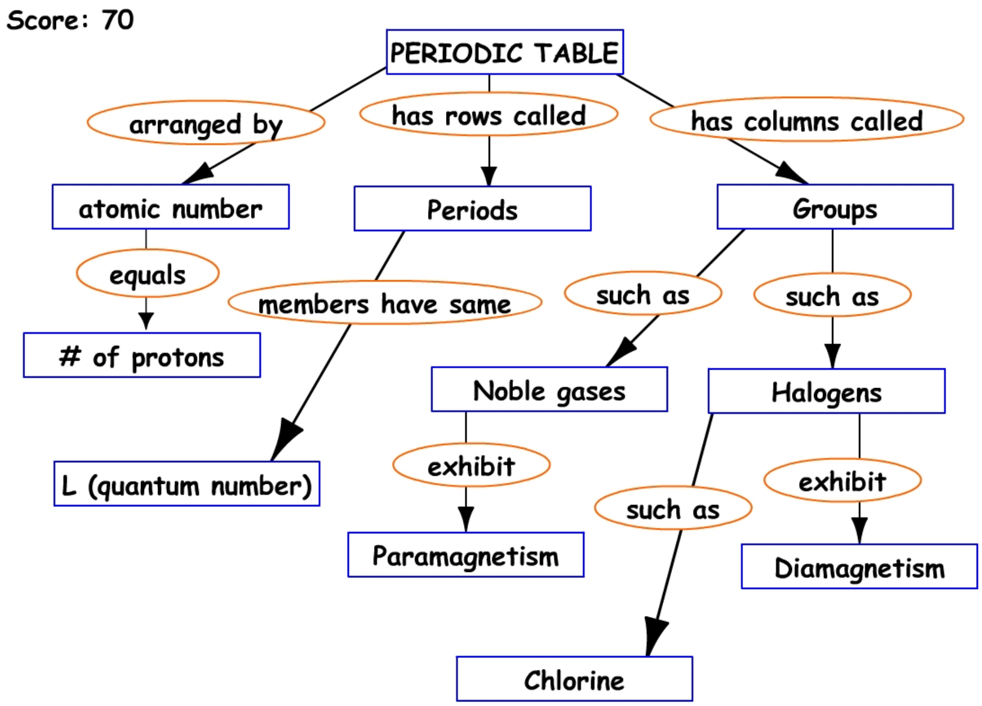
Student 29

No concept map submitted

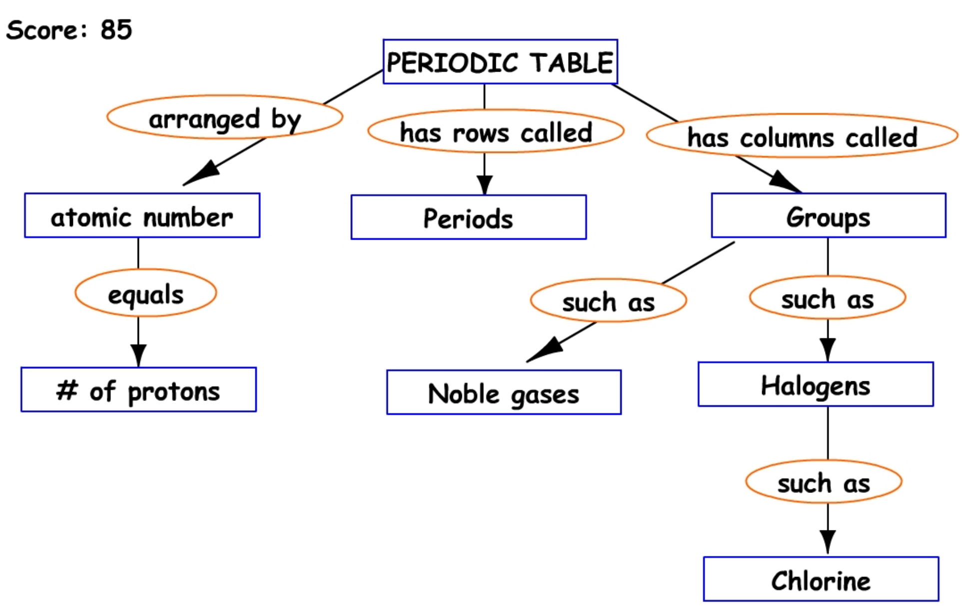
Student 30



Student 31



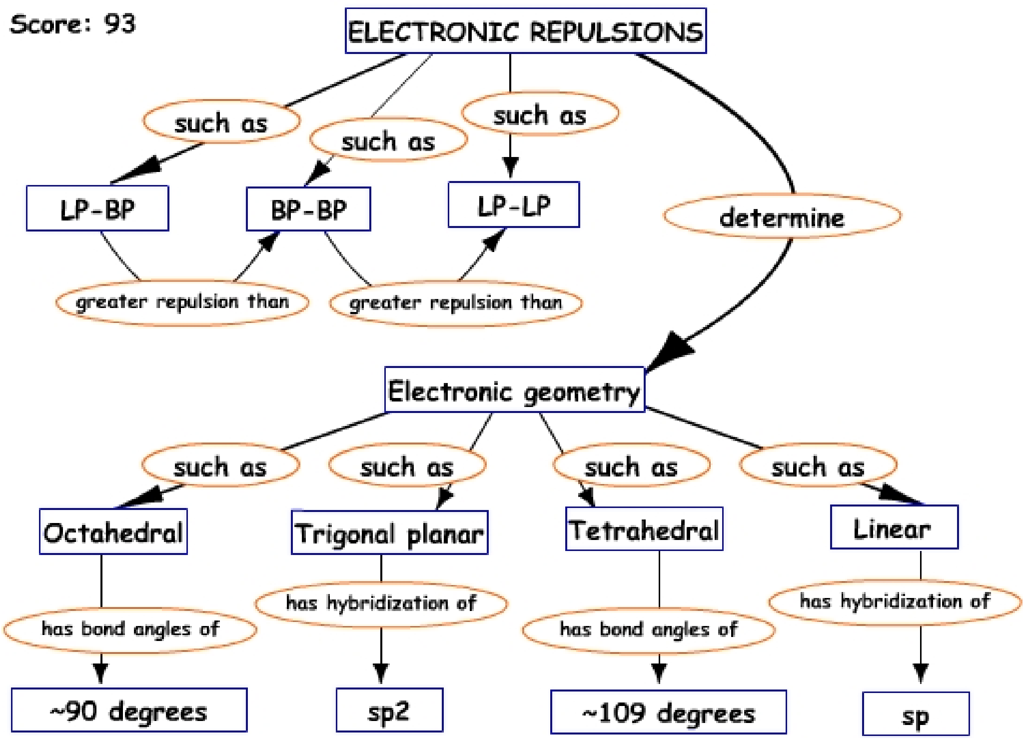
Student 32



APPENDIX C. CONCEPT MAPPING ASSIGNMENT 2

Concept maps submitted by students for Concept Mapping Assignment 2 were drawn in CMAT and stored as a list of propositions in a text file. The submissions for the 32 students enrolled in the CH 301wb course were redrawn into concept maps used for the hand grading procedures in this study. Students who did not submit a concept maps are noted as "*No concept map submitted*."

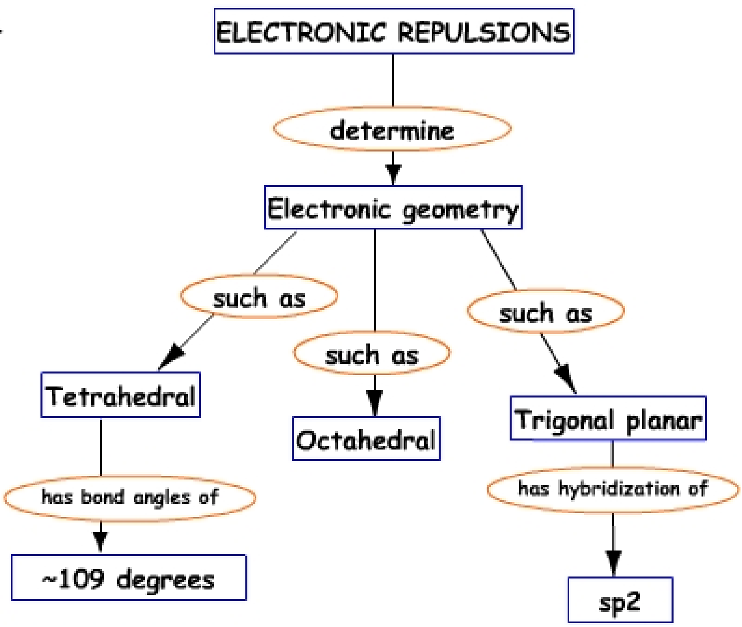
Student 1



Student 2
No concept map submitted

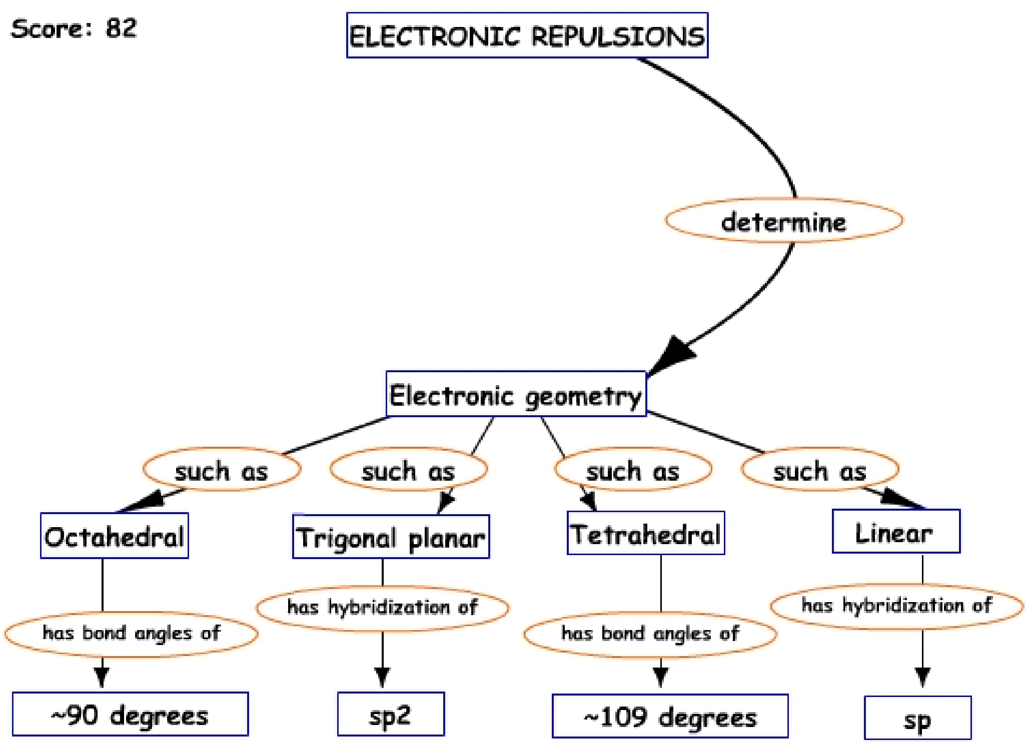
Student 3

Score: 71



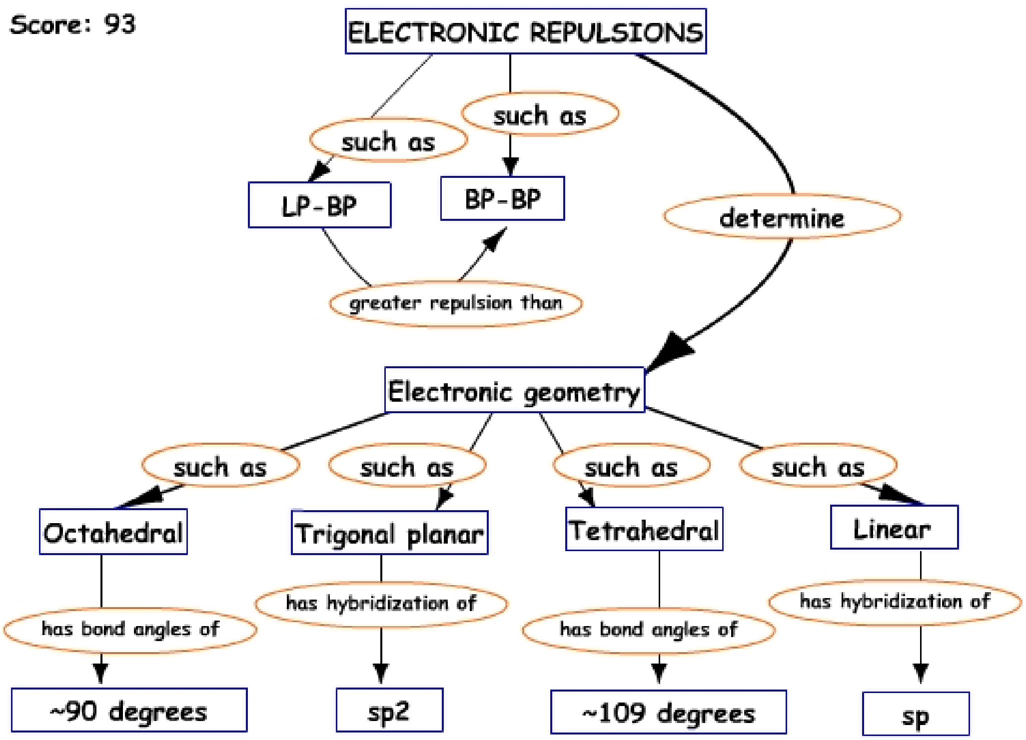
Student 4

Score: 82



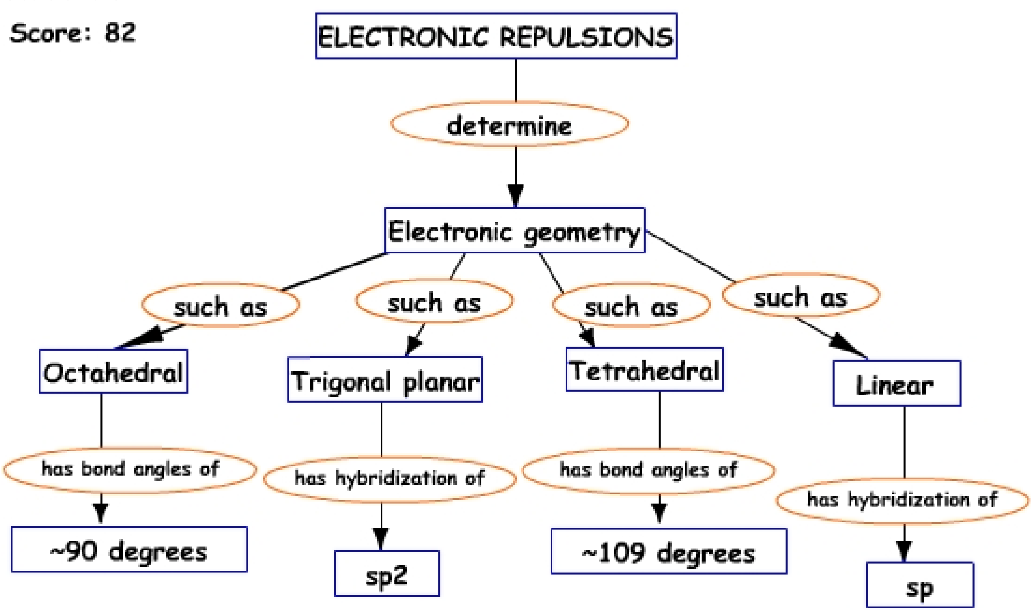
Student 5

Score: 93



Student 6

Score: 82

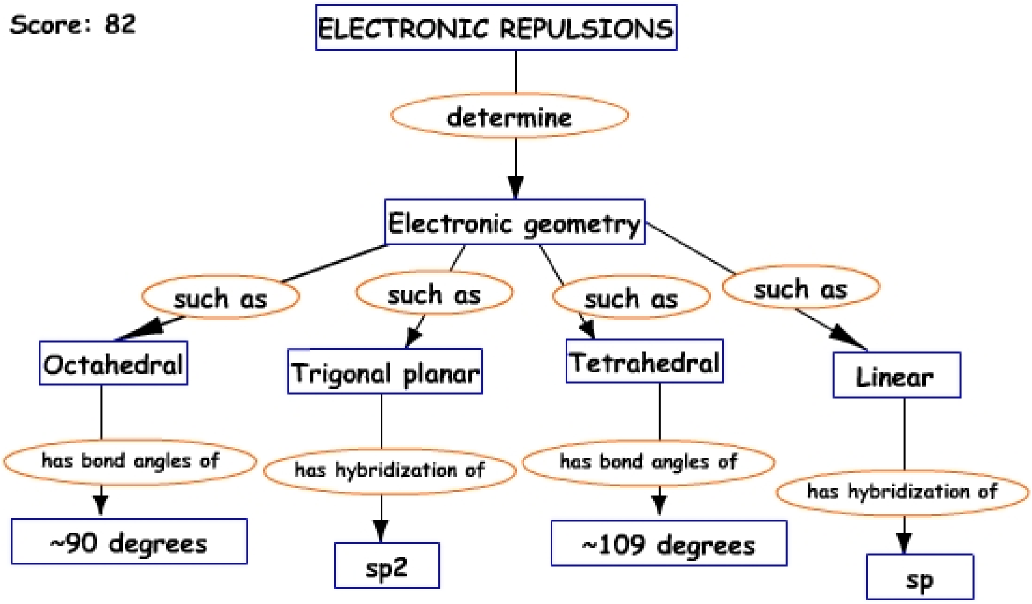


Student 7

No concept map submitted

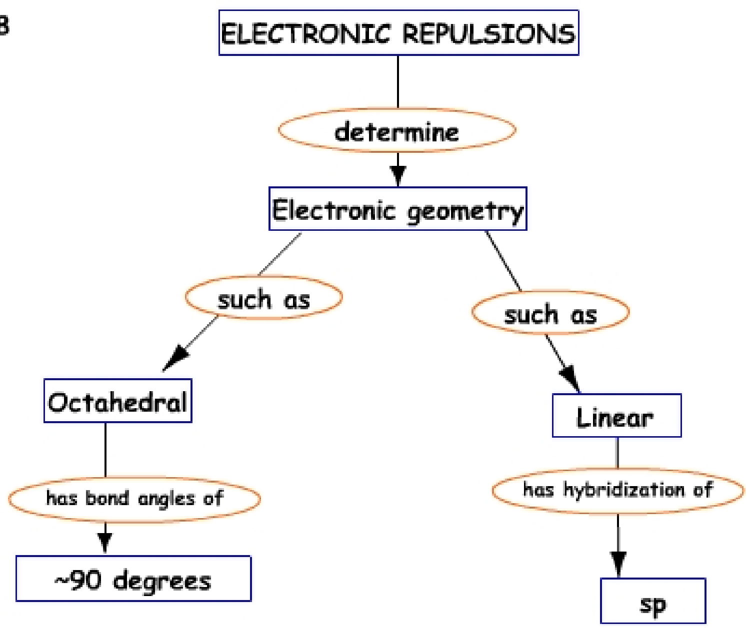
Student 8

Score: 82

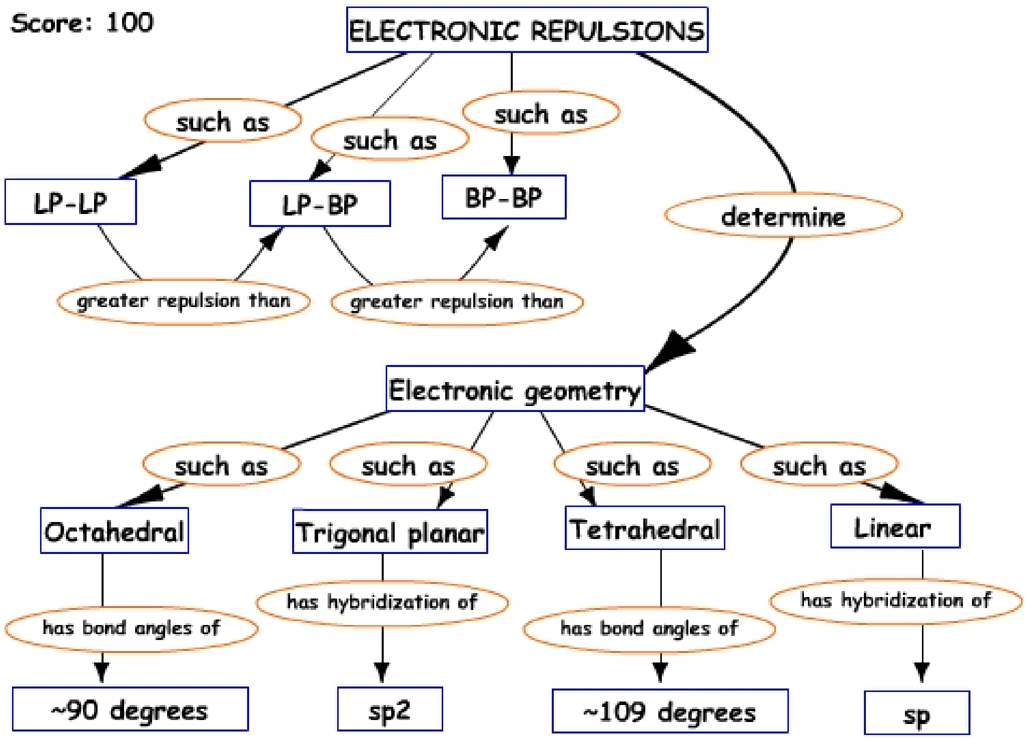


Student 9

Score: 68



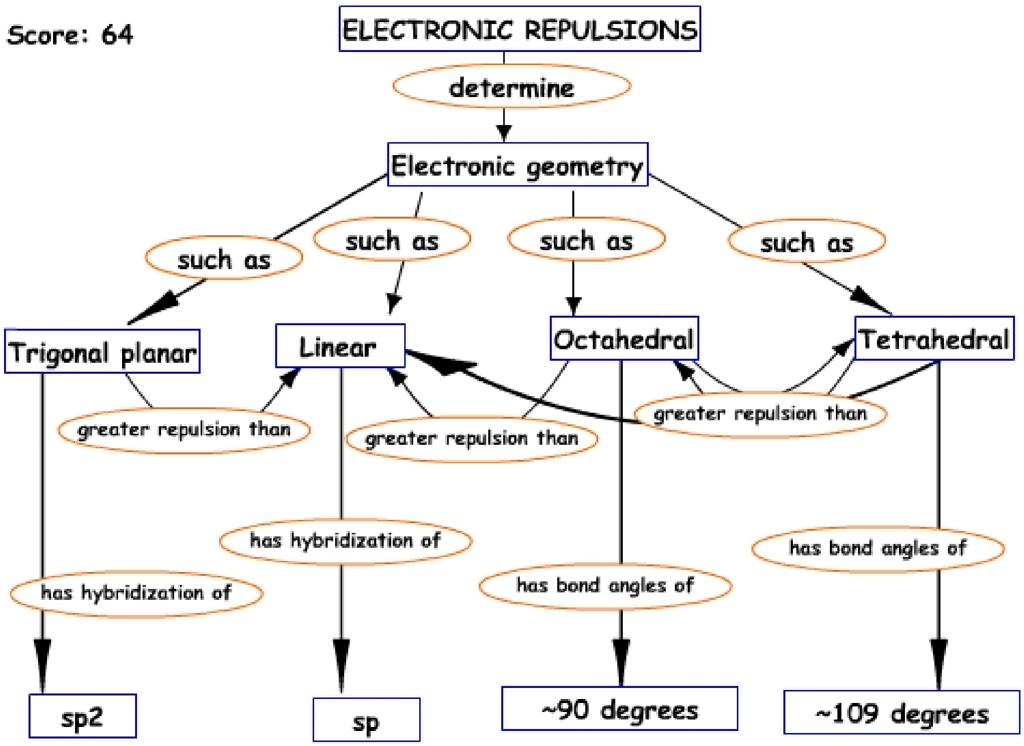
Student 10



Student 11

No concept map submitted

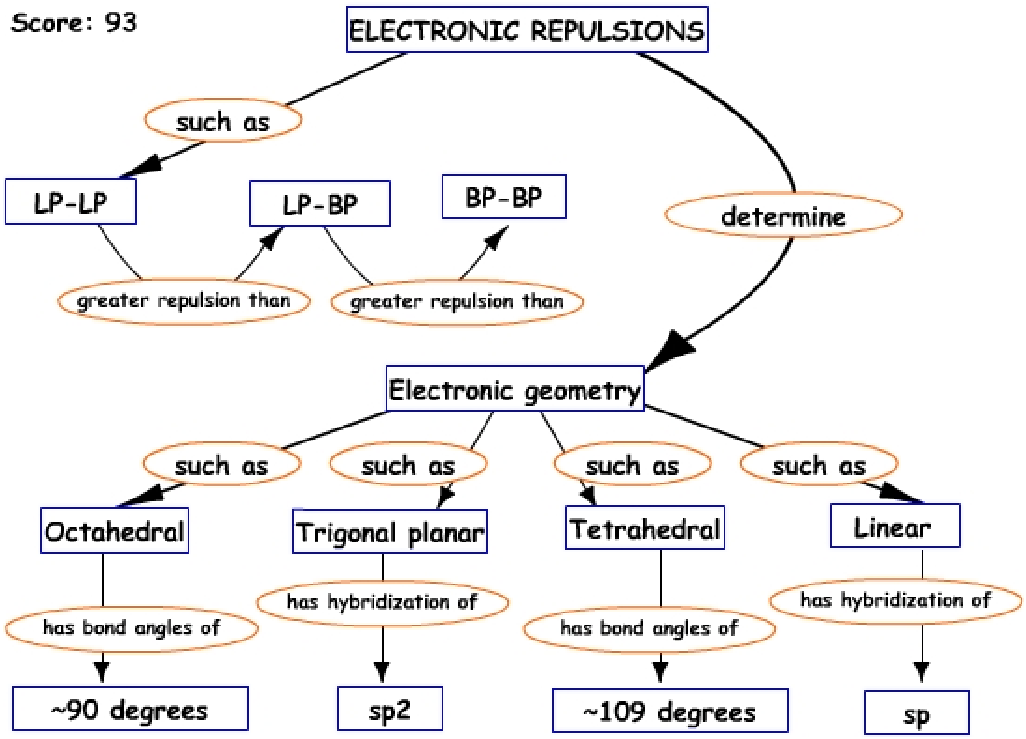
Student 12



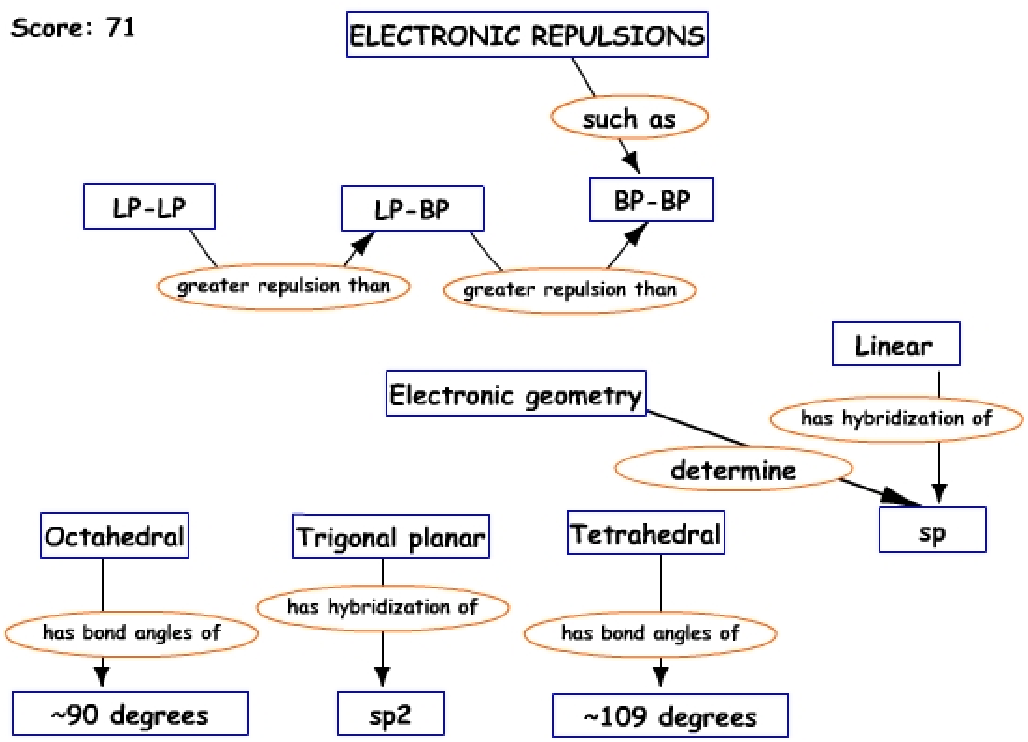
Student 13

No concept map submitted

Student 14

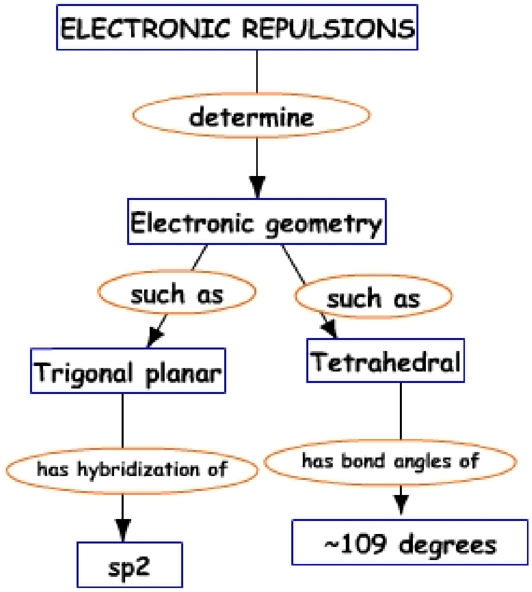


Student 15



Student 16

Score: 68



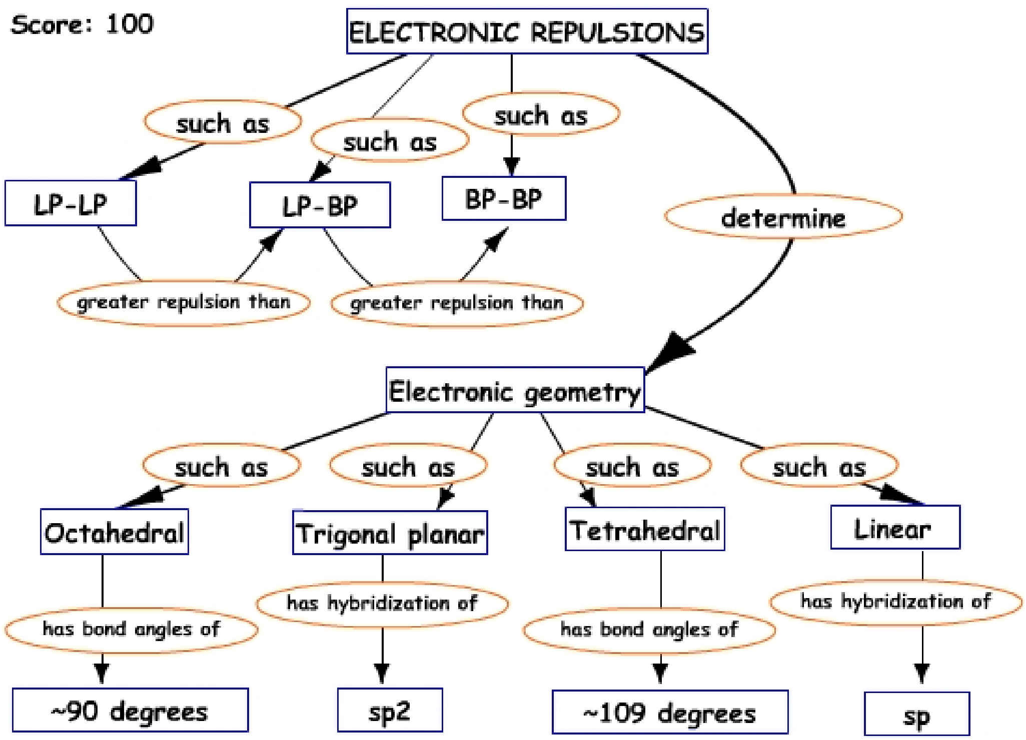
Student 17

No concept map submitted

Student 18

No concept map submitted

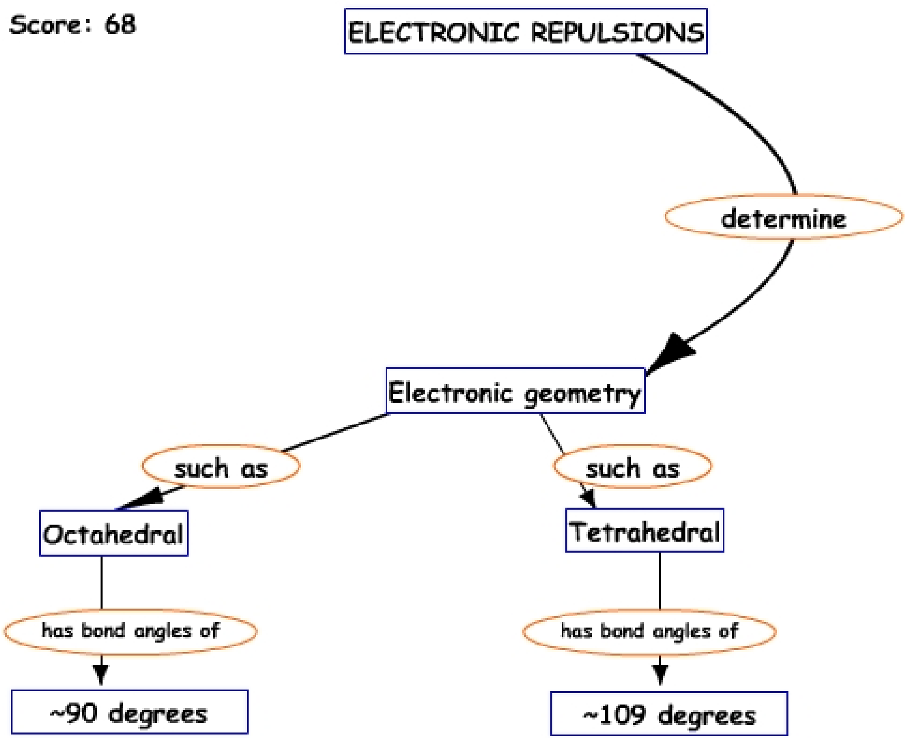
Student 19



Student 20

No concept map submitted

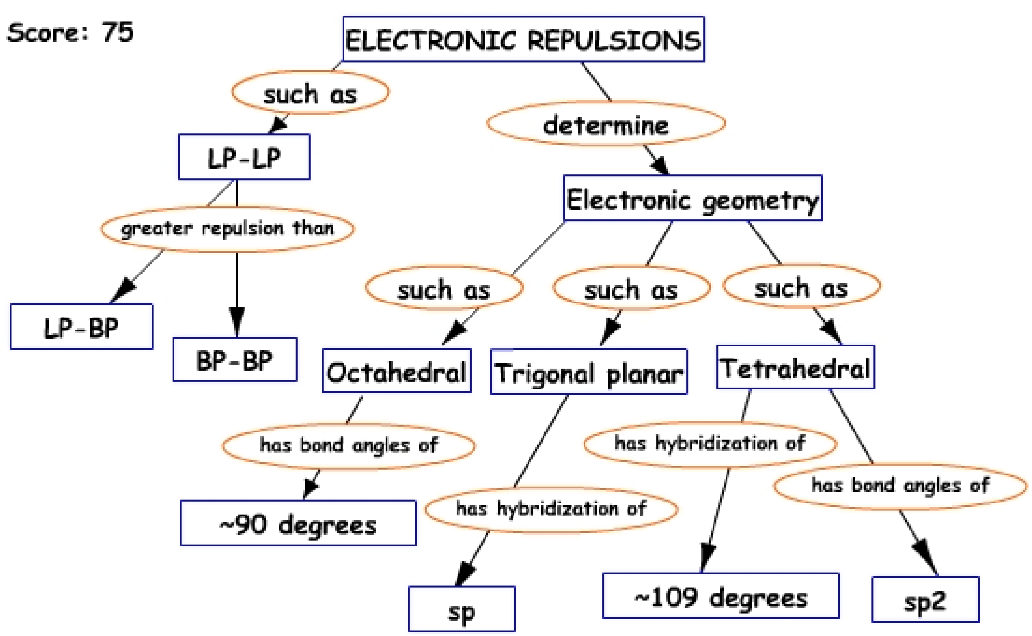
Student 21



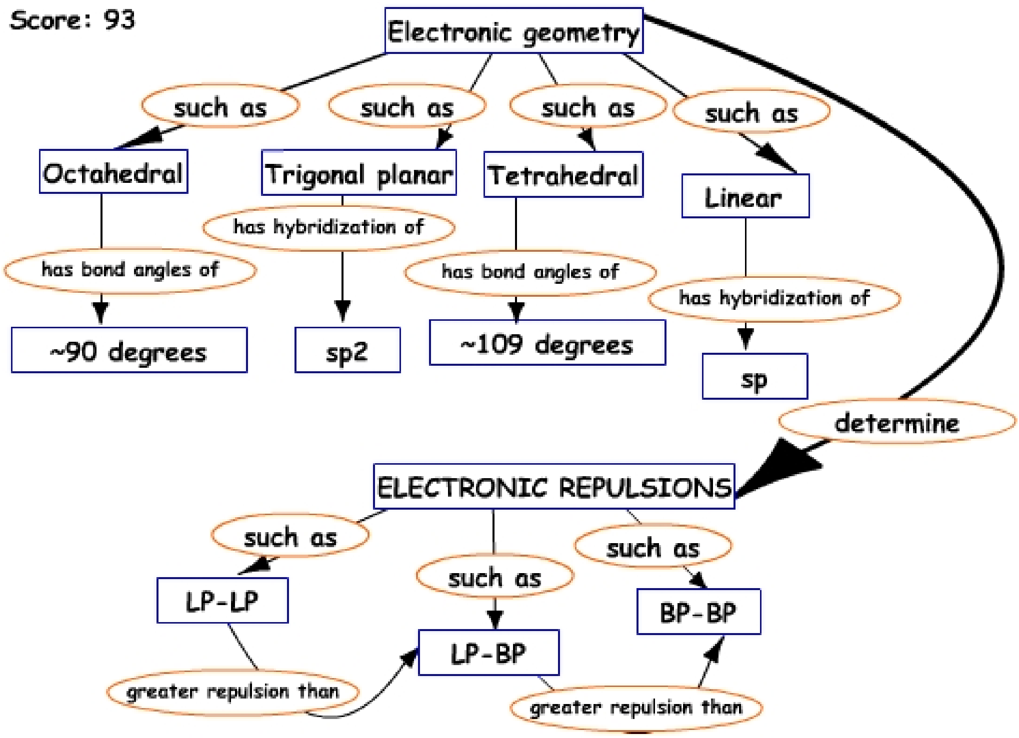
Student 22

No concept map submitted

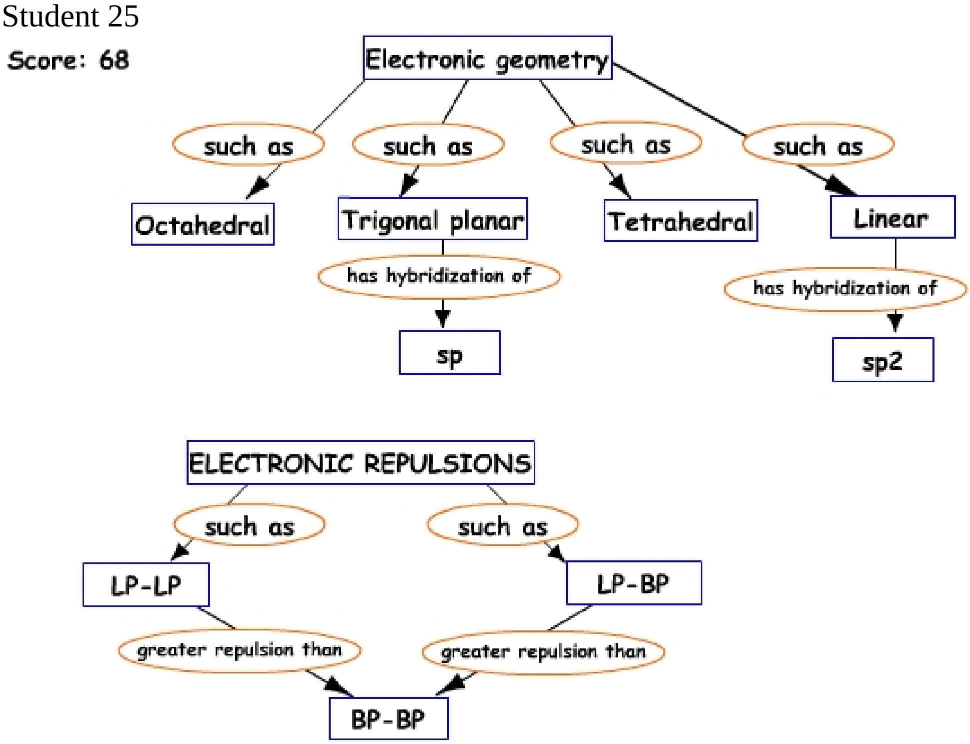
Student 23



Student 24



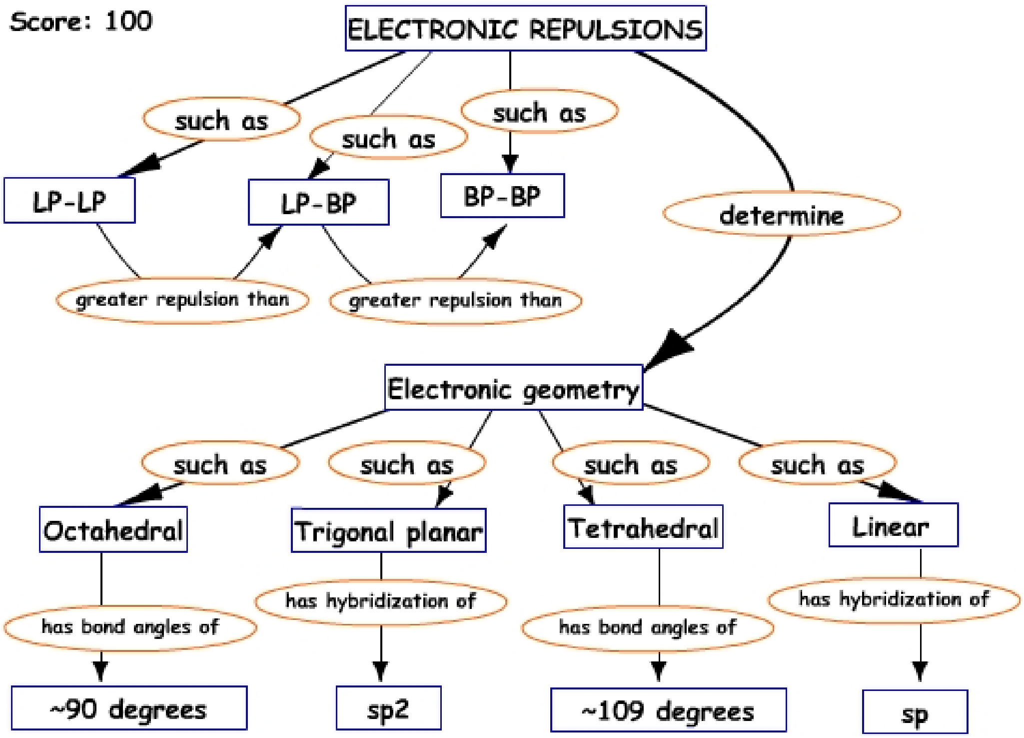
Student 25



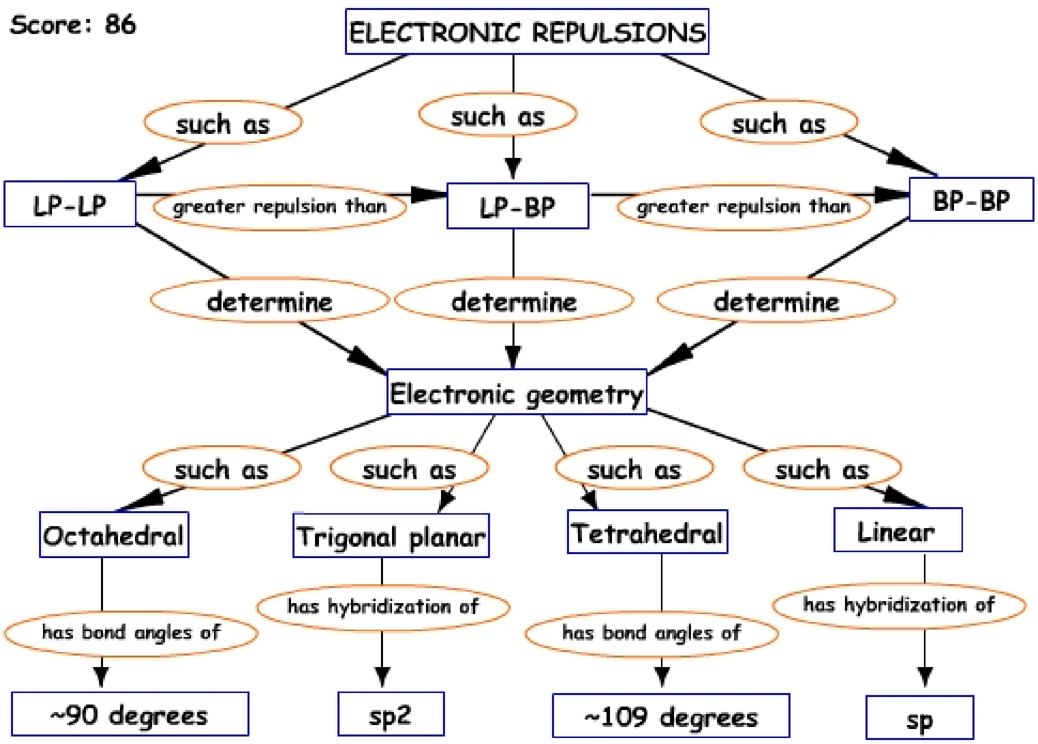
Student 26

No concept map submitted

Student 27



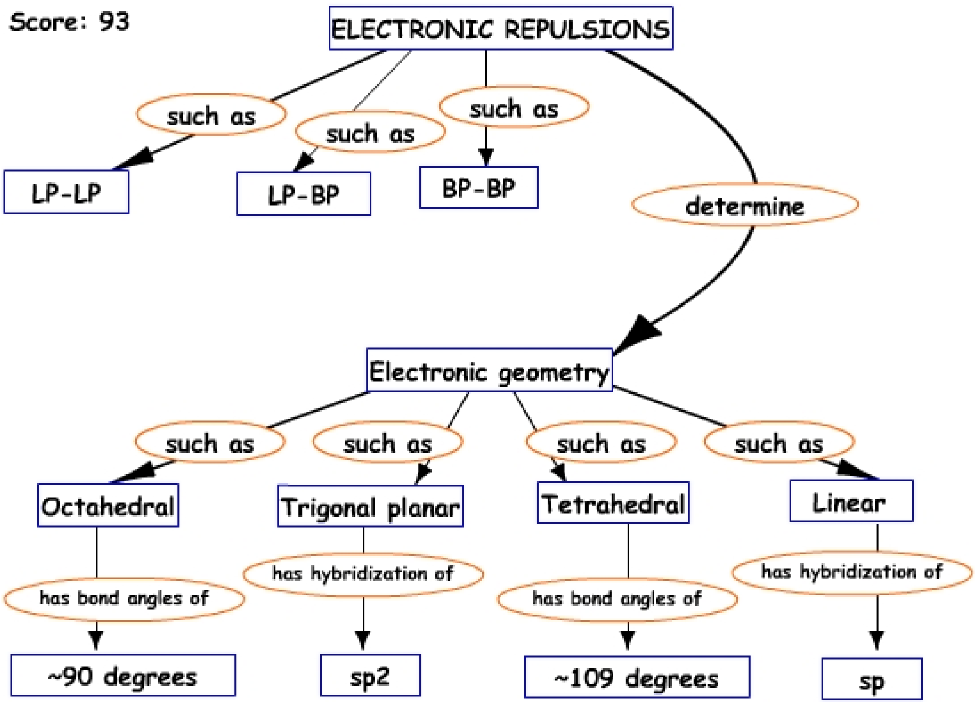
Student 28



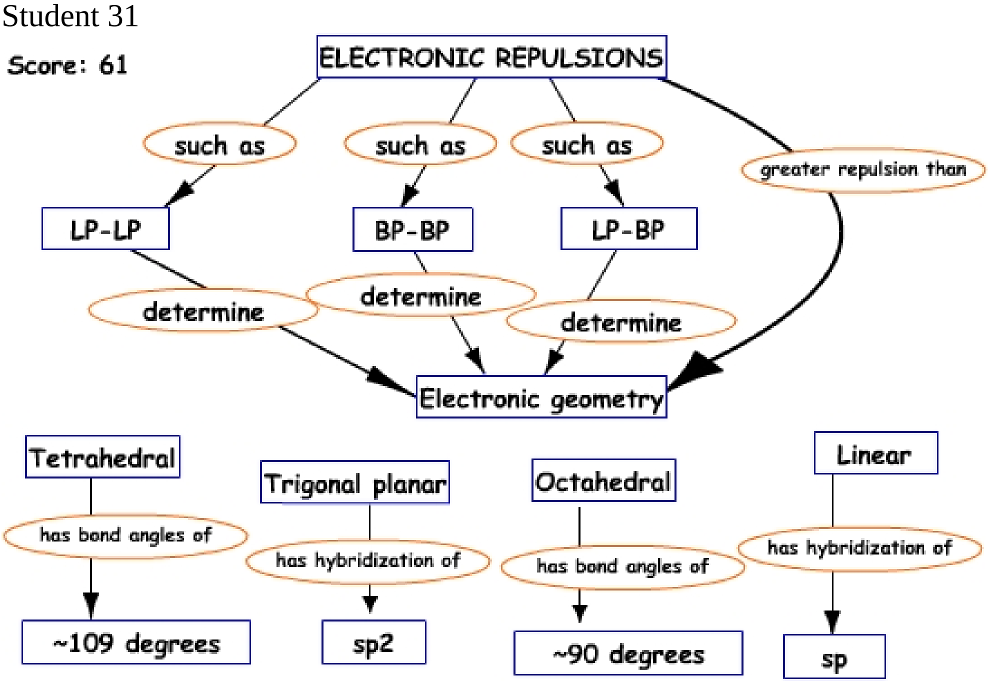
Student 29

No concept map submitted

Student 30



Student 31



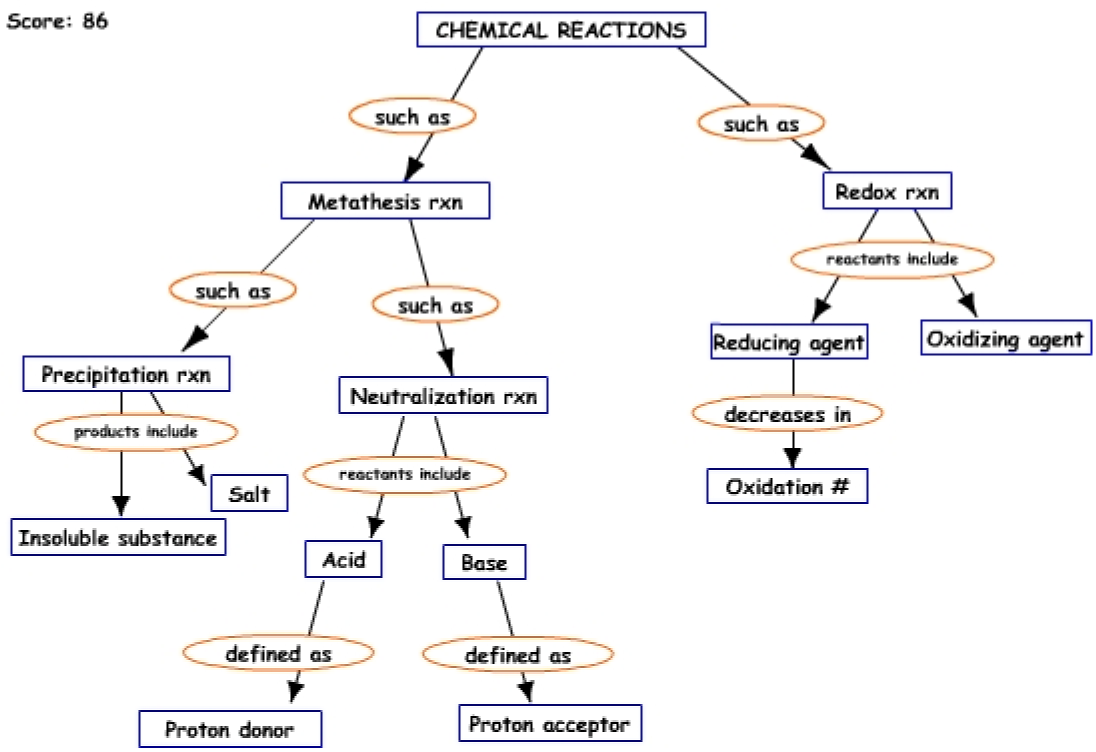
Student 32

No concept map submitted

APPENDIX D. CONCEPT MAPPING ASSIGNMENT 3

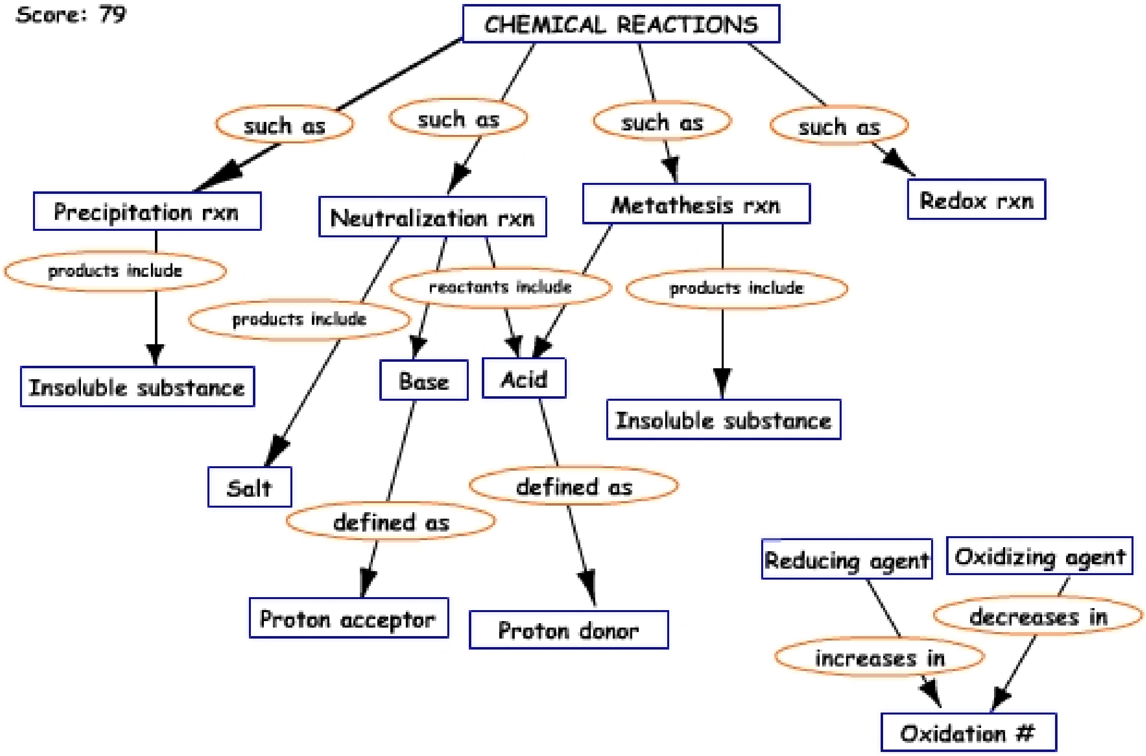
Concept maps submitted by students for Concept Mapping Assignment 3 were drawn in CMAT and stored as a list of propositions in a text file. The submissions for the 32 students enrolled in the CH 301wb course were redrawn into concept maps used for the hand grading procedures in this study. Students who did not submit a concept maps are noted as "*No concept map submitted*."

Student 1



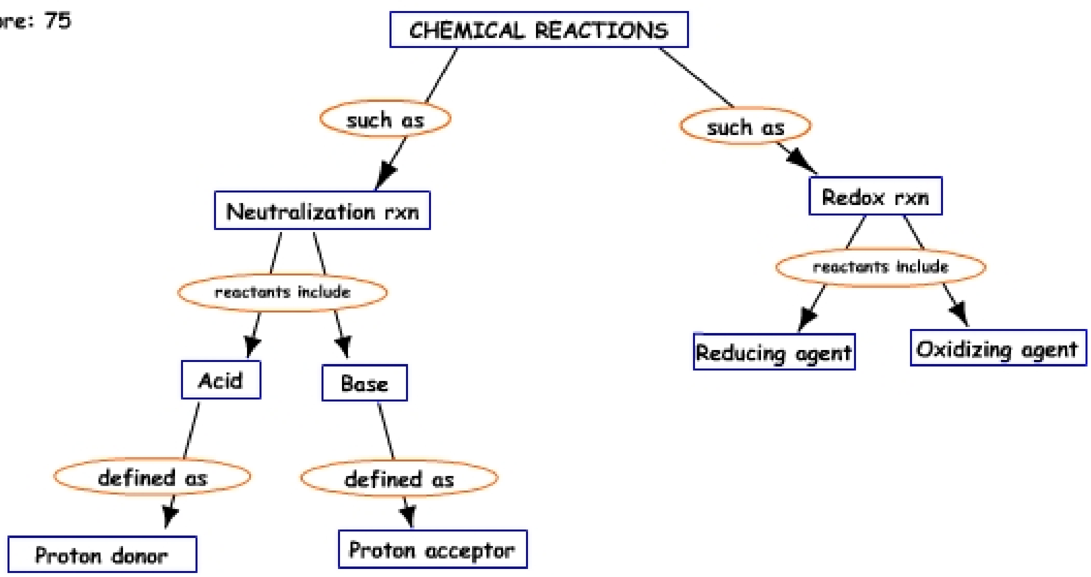
Student 2

Score: 79



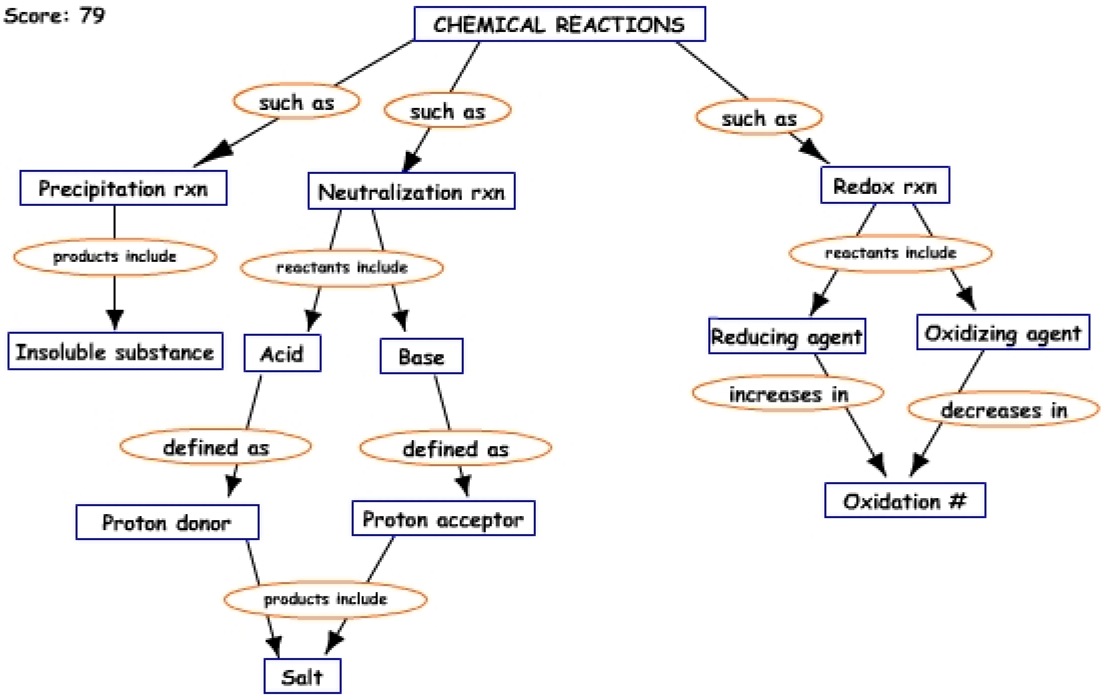
Student 3

Score: 75



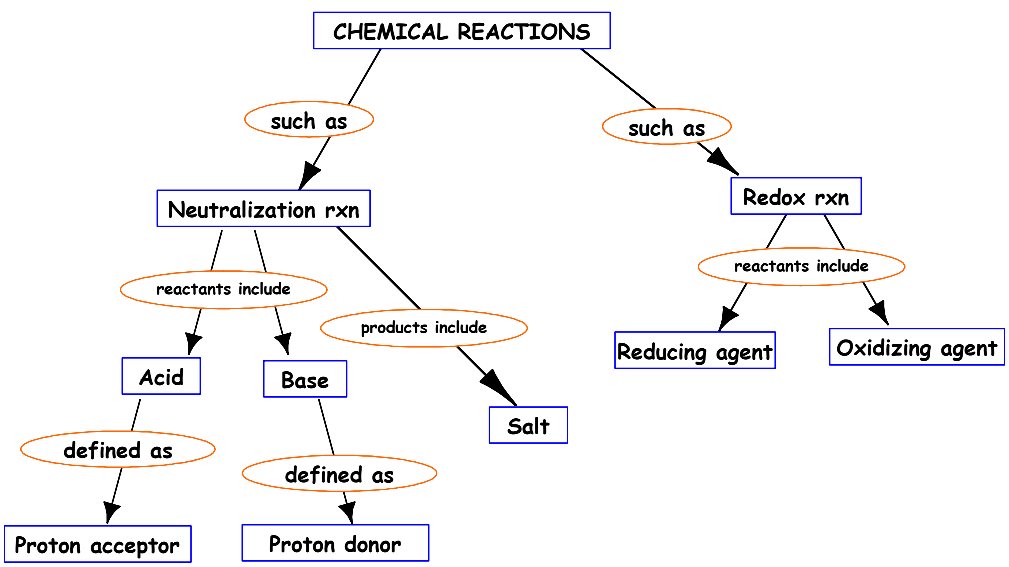
Student 4

Score: 79



Student 5

Score: 64

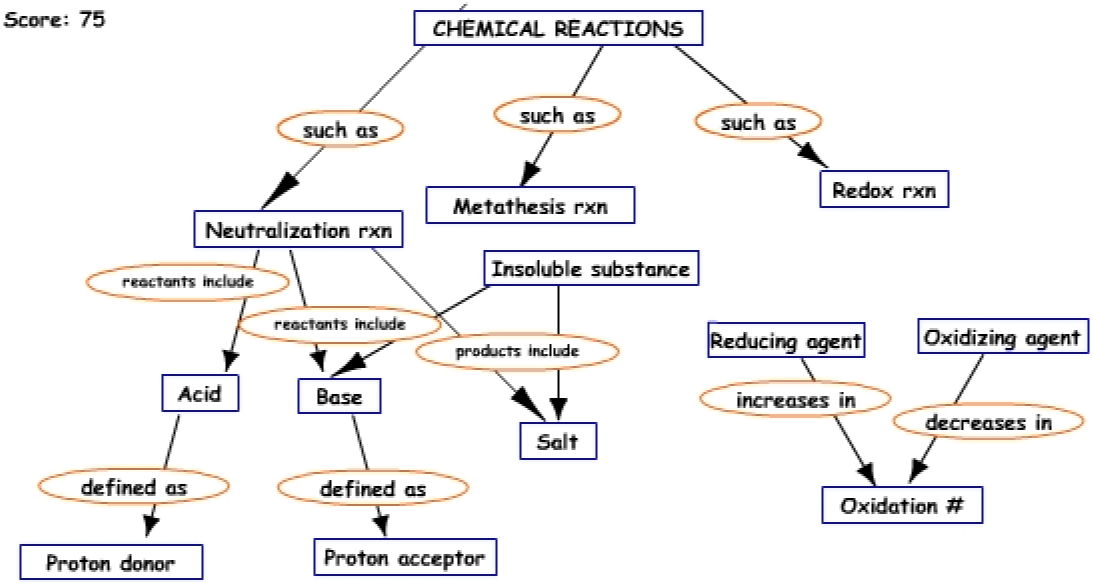


Student 6

No concept map submitted

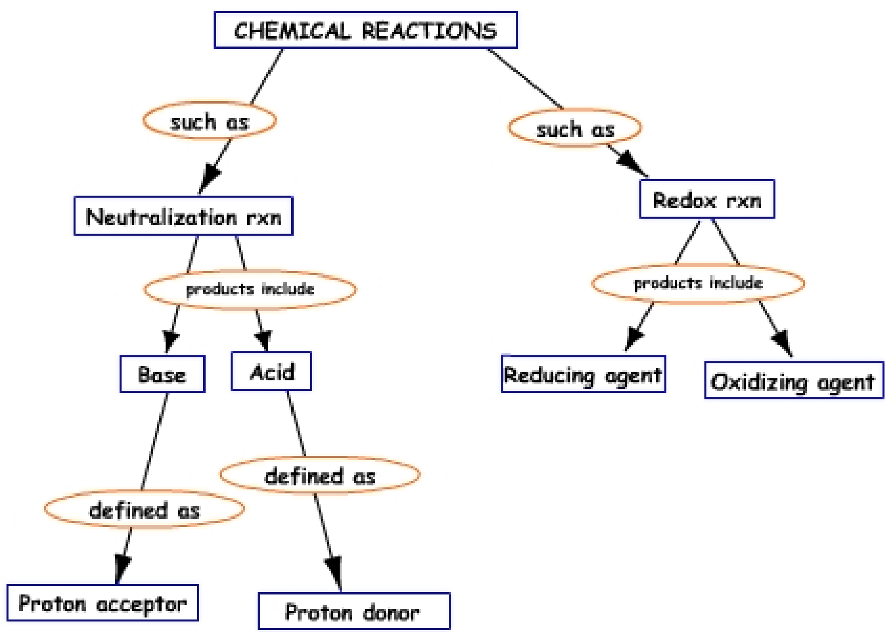
Student 7

Score: 75



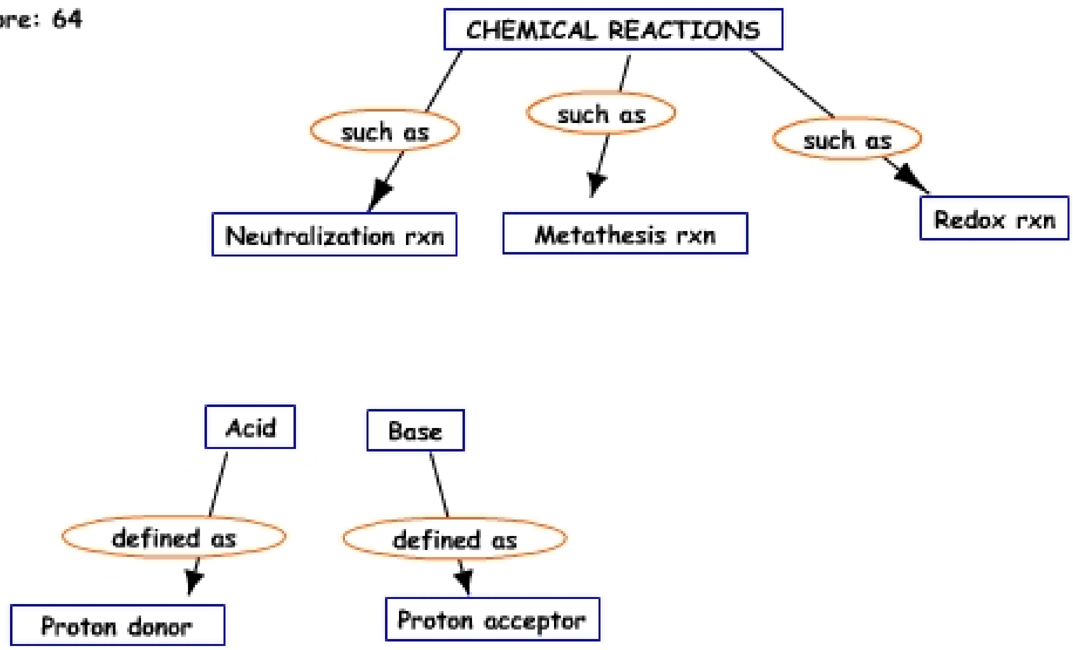
Student 8

Score: 54



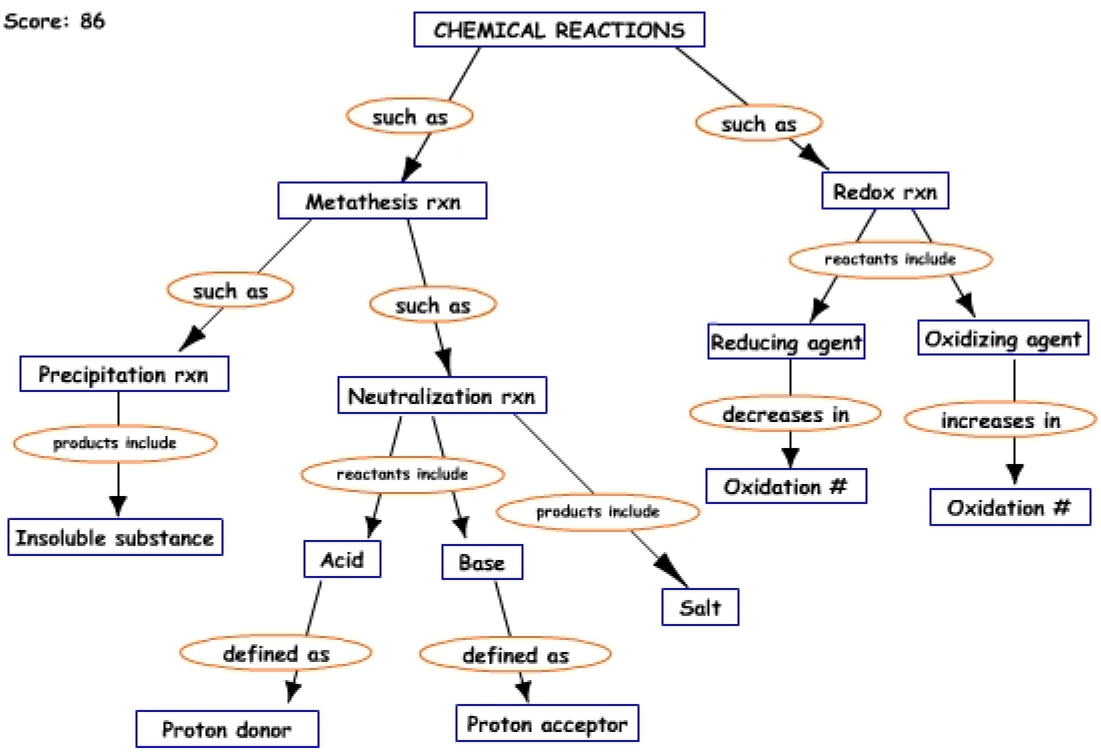
Student 9

Score: 64



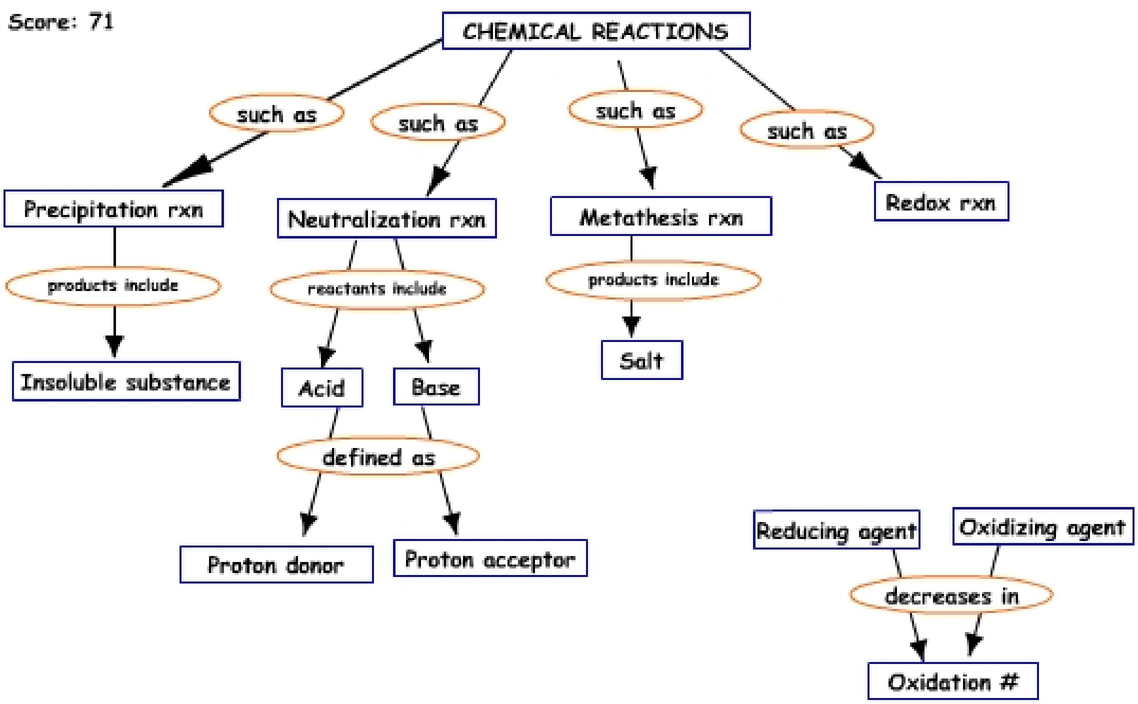
Student 10

Score: 86

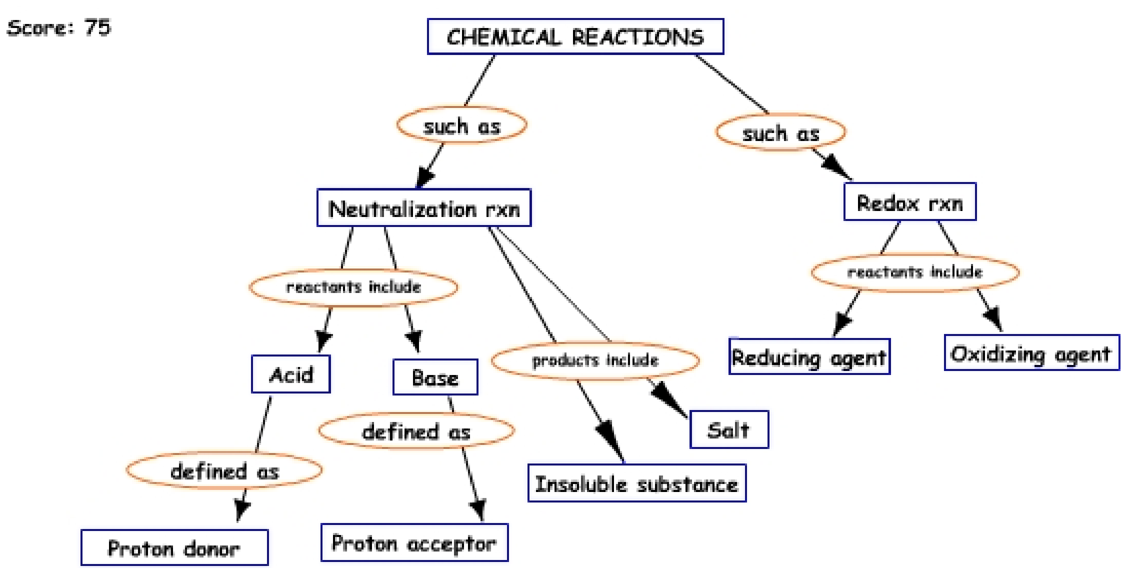


Student 11
No concept map submitted

Student 12



Student 13

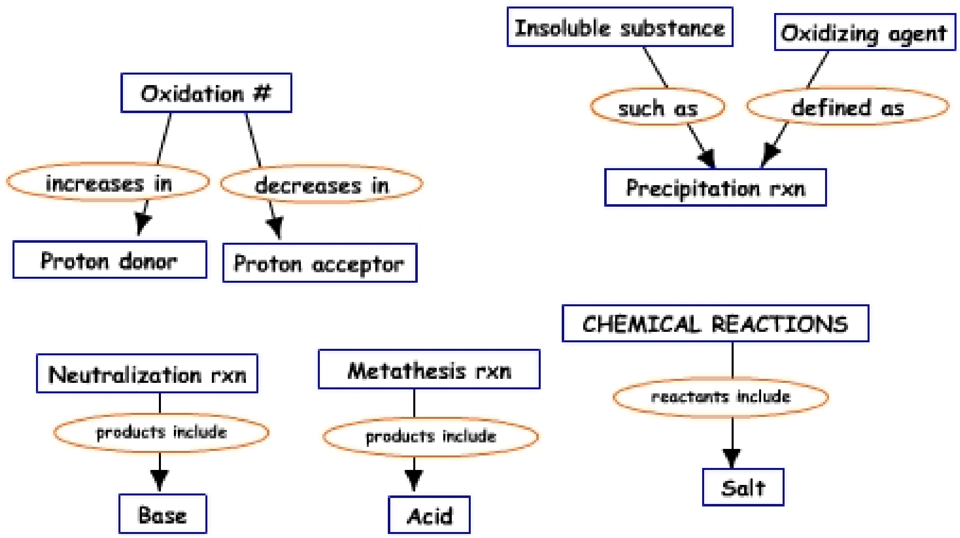


Student 14

No concept map submitted

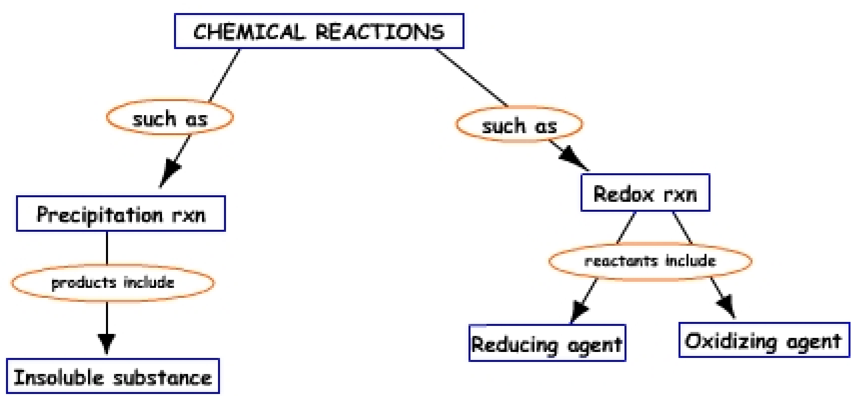
Student 15

Score: 29



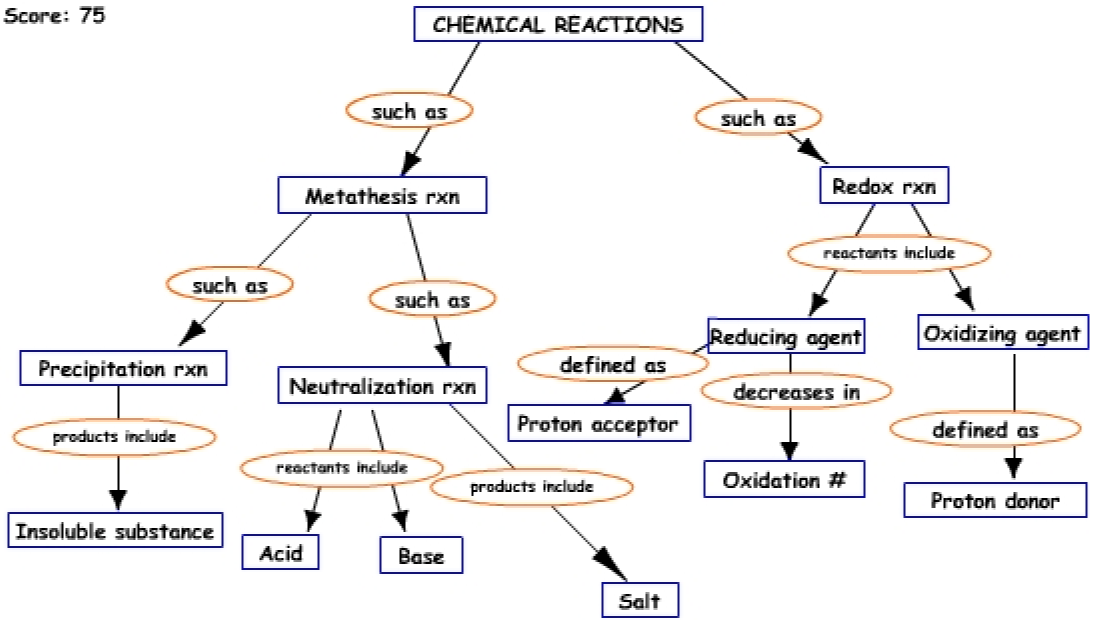
Student 16

Score: 64



Student 17

Score: 75

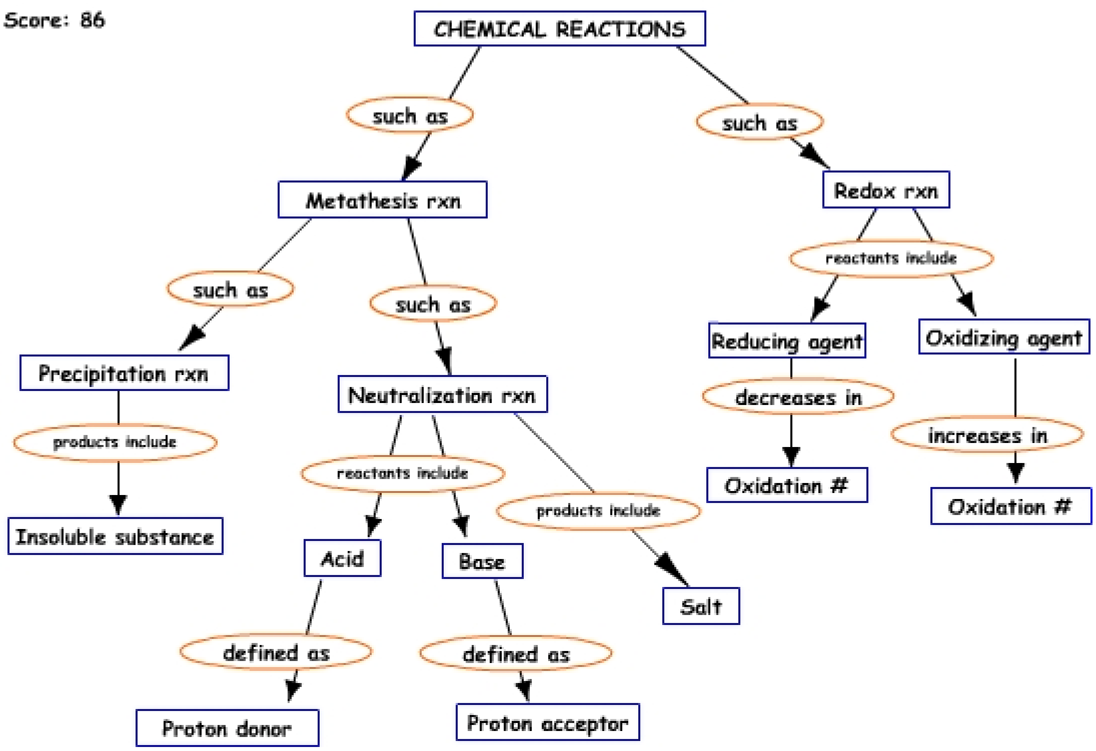


Student 18

No concept map submitted

Student 19

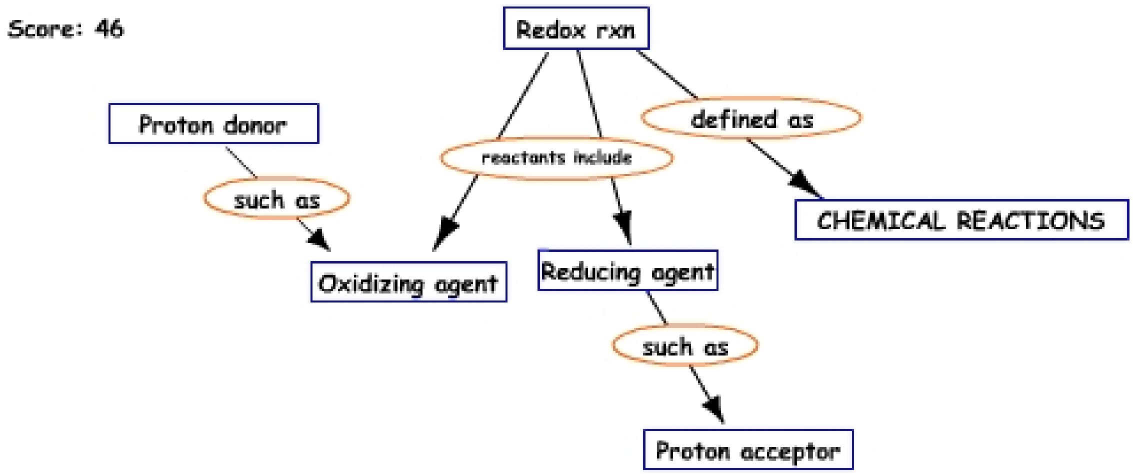
Score: 86



Student 20

No concept map submitted

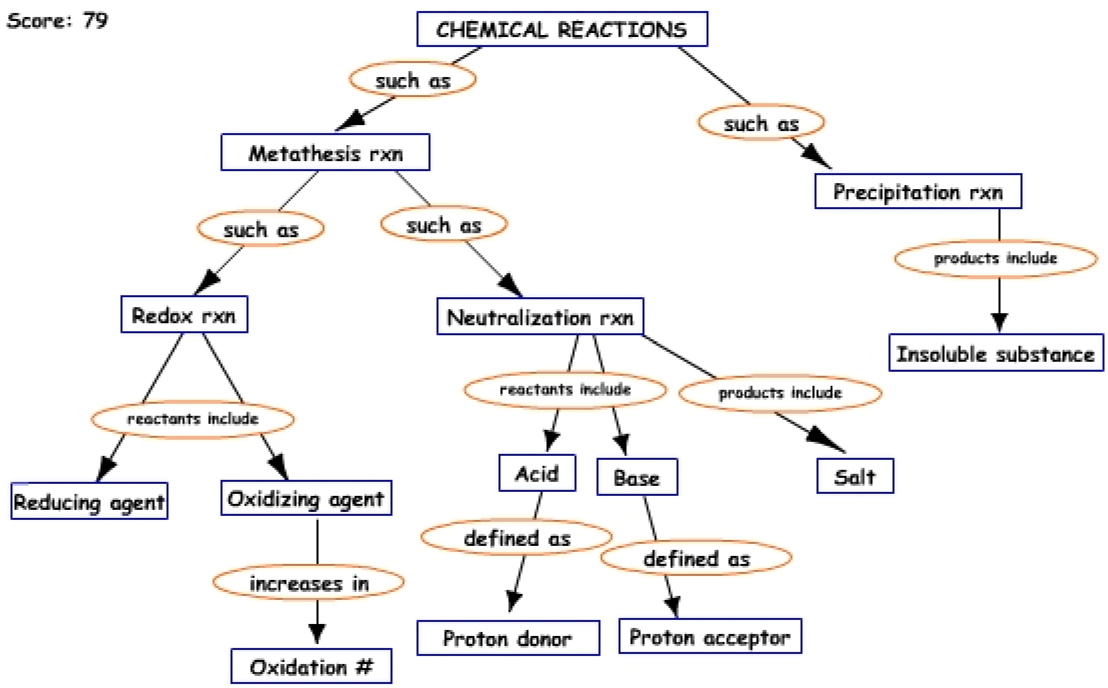
Student 21



Student 22

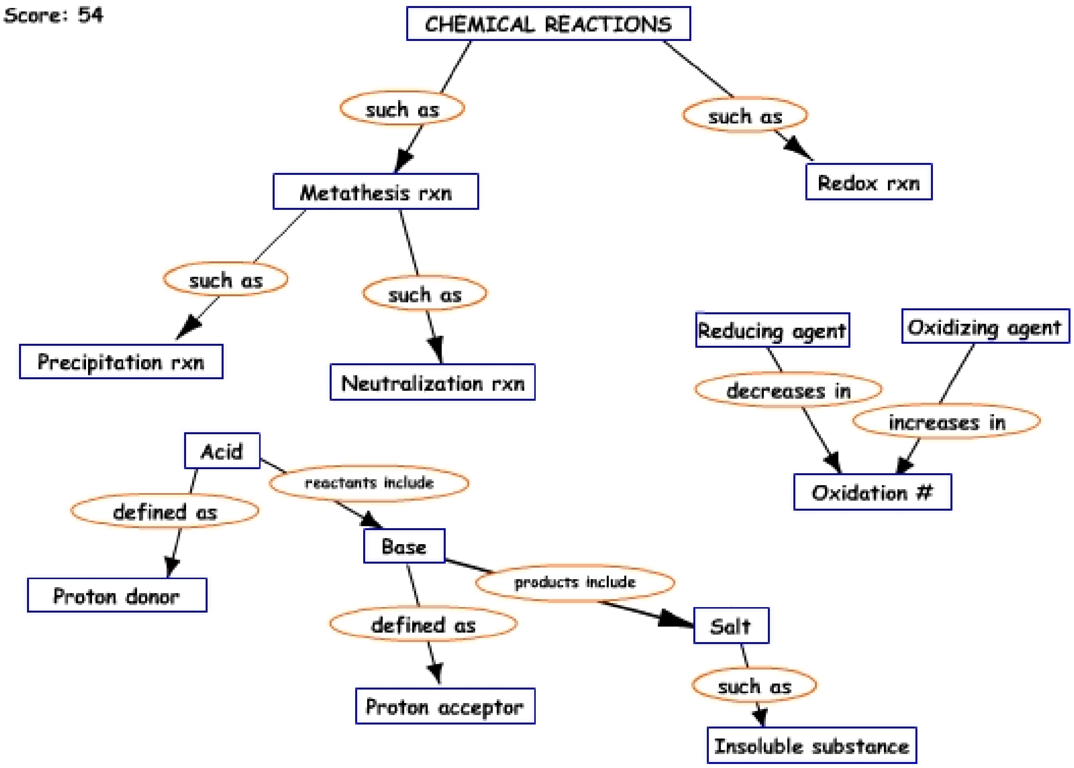
No concept map submitted

Student 23



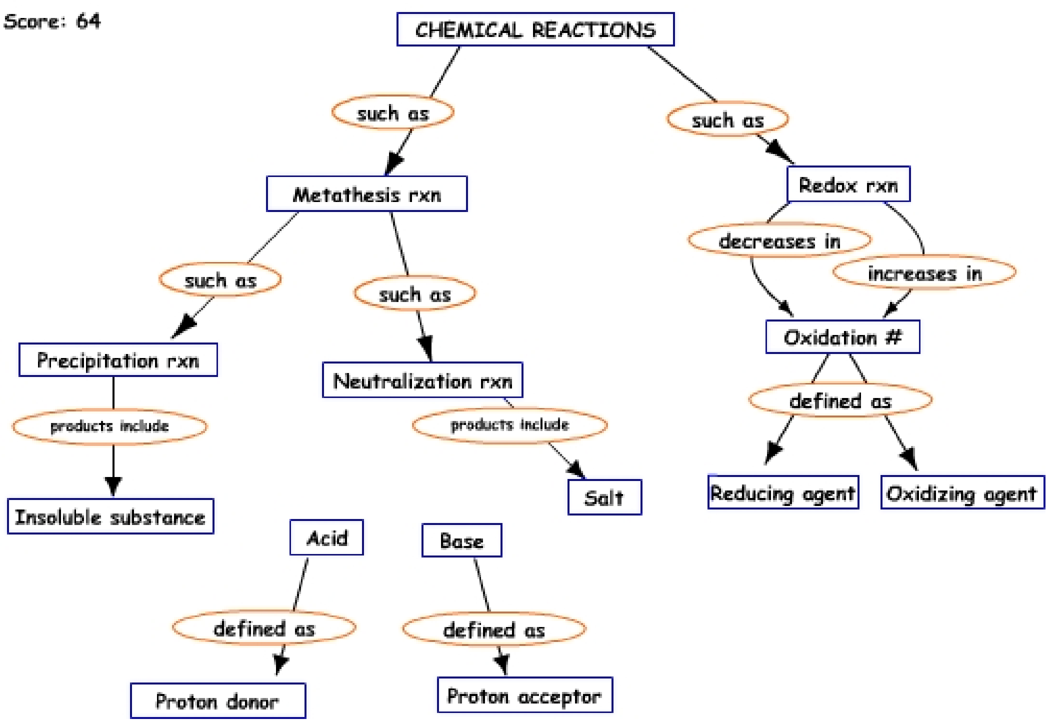
Student 24

Score: 54



Student 25

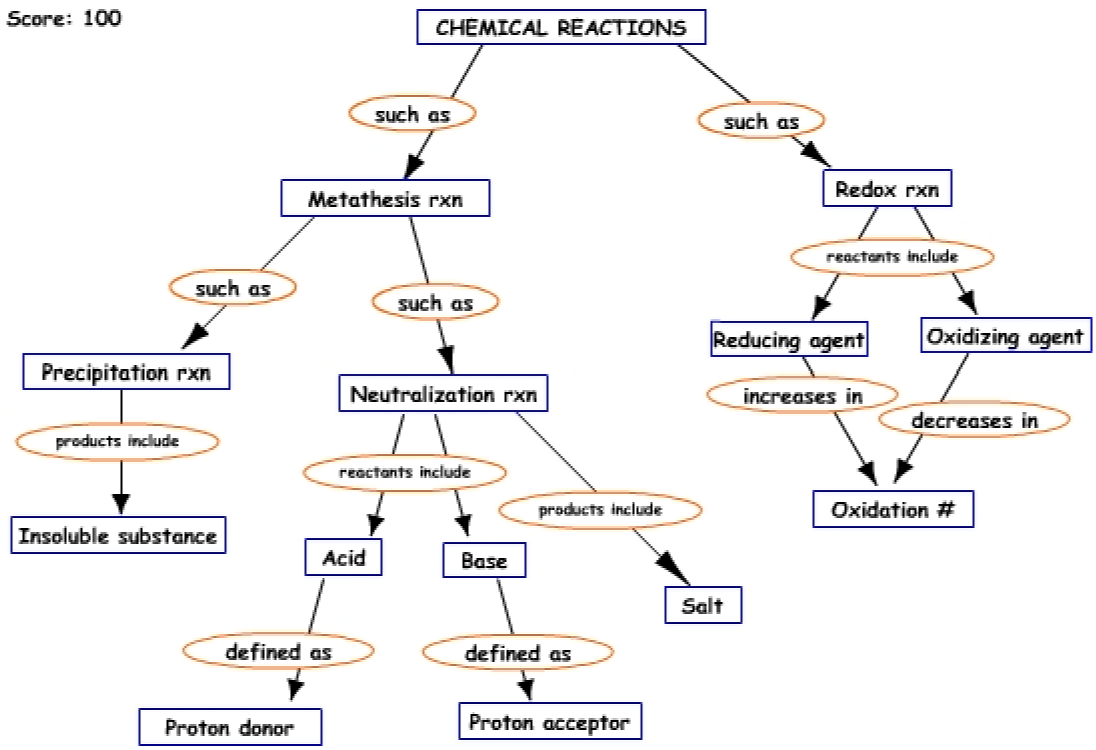
Score: 64



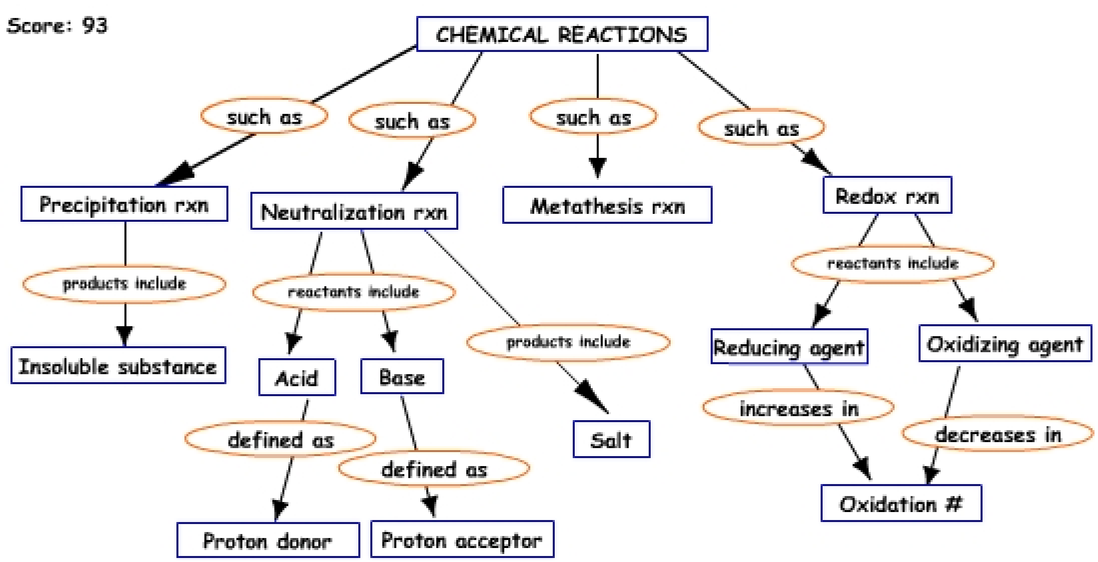
Student 26

No concept map submitted

Student 27

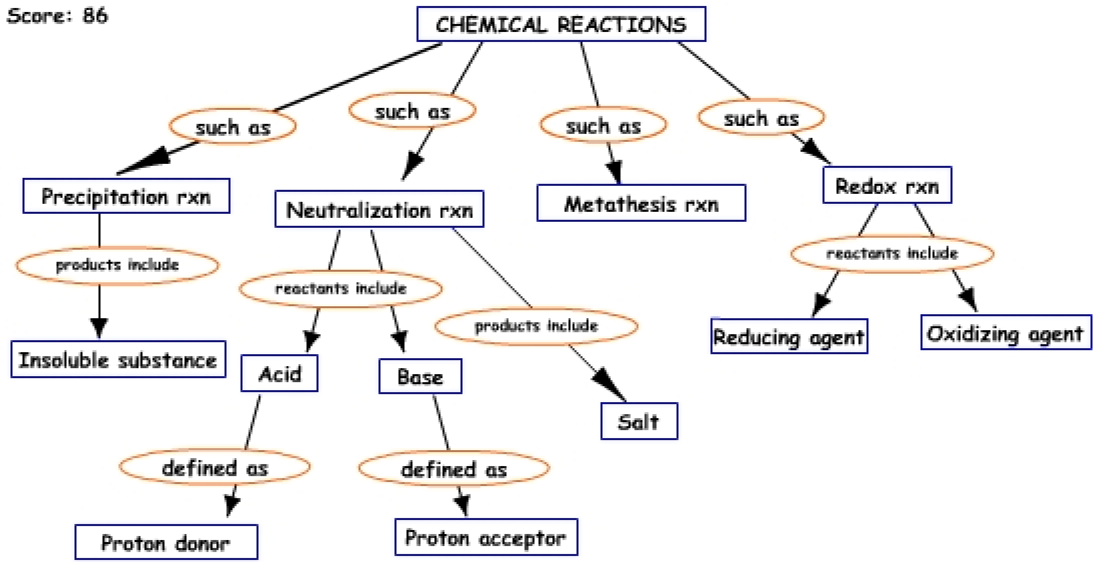


Student 28



Student 29
No concept map submitted

Student 30



Student 31
No concept map submitted

Student 32
No concept map submitted

Glossary

Assessment

The process of acquiring information in measurable terms, knowledge, or skills used to make an evaluation of learning.

Assessment tool

The activity used for the purpose of assessment, often called an instrument. Typical assessment tools are tests and questionnaires, or concept maps.

Bloom's Taxonomy

Bloom headed a commission in the 1950's which set out to describe the three domains in the educational process: the cognitive domain, the affective domain, and the psychomotor domain. The taxonomy referred to in this study is the cognitive domain. The Taxonomy of the Cognitive Domain is comprised of six levels arranged in ascending order of difficulty are:

- **Knowledge**: Recognition and recall of previously learned facts or basic concepts.
- **Understanding**: Interprets, translates or summarizes given information in own terms.
- **Application**: Uses information in a new situation to solve problems.
- **Analysis**: Separates complex concepts into more simple concepts until relationships are clear.
- **Synthesis**: Combines elements to create a product with new meaning or purpose.
- **Evaluation**: Judging or decision–making based on criteria or rationale.

Concept

An abstract notion which places a label representing a mental picture of a group of things that have common characteristics, concepts represent physical objects, such as a desk or a chair; or abstract ideas, such as ‘conservation of mass’ and a mole.

Concept Map

An arrangement of major concepts into a visual arrangement formed by connecting related concepts with labeled directional lines. Concept maps reveal the structural pattern for the purpose to form or assess a person's knowledge structure.

Constructivism

Constructivism is a theory of learning whose practitioners subscribe to the notion that learning is an individual process where previous experiences affect the manner in which new information is viewed and eventually learned. Cognitive scientists like Piaget and Ausubel have been linked to this school of learning by describing the acquisition of new concepts being based upon previous knowledge.

Formative Assessment

Type of assessment used to aid learning by providing feedback on a student's work, and would not necessarily be used for grading purposes.

Meaningful Learning

New knowledge is anchored to previously existing knowledge in a non-arbitrary, non-verbatim manner. Meaningful learning is contrasted to rote learning where new knowledge is related to existing knowledge in a verbatim manner (memorization).

Misconception

Referred to as an “alternative conception” by some, it is a relationship between two concepts a learner forms in their cognitive structure that is different or unaccepted as the standard or correct relationship.

Scoring (concept maps)

Scoring is the process of assigning value to individual items in an assignment.

- Relational scoring: Scoring a proposition on the basis of its factual correctness
- Structural scoring: Scoring a concept map based upon its structural characteristics, such as branching points, number of chains formed, and chain length

Summative Assessment

Type of assessment typically carried out at the end of a course to assign students a course grade.

References

1. Moore, J. W. (1989). *Journal of Chemical Education, 66,* 15.
2. Hoffman, B. (1962). *The Tyranny of Testing.* The Crowell-Collier Press.
3. Bowen, C. W. and Phelps, A. J. (1997) *Journal of Chemical Education, 74,* 715.
4. Zoller, U.; Fastow, M.; Lubezky, A., and Tsaparlis, G. (1999) *Journal of Chemical Education, 76,* 112.
5. Phye, G. D. (1997) *Handbook of Academic Learning: Construction of knowledge.* Academic Press, San Diego, CA.
6. National Research Council (NRC) (1996). *National Science Education Standards.* Washington D.C.: National Academy Press.
7. Hobson, E.H. (1997). *The Clearing House, 71* (2), 68.
8. Holmes, G.A. and Leitzel, T.C. (1993). *Performance & Instruction, 32* (8), 28.
9. Bailey, P.D. and Garratt, J. (2002). *University Chemical Education, 6,* 39.
10. Web test tools website, http://www.webtesttools.com, accessed August 1, 2005.
11. Bloom, B. S. ed. (1956) *Taxonomy of Educational Objectives, Handbook 1: The Cognitive Domain*. New York: David McKay Co. Inc.
12. Horton, P.B., *et. al.* (1993). *Science Education, 77*, 95.
13. Kinchin, I. M. (2001). *International Journal of Science Education, 23*, 1257.
14. Ruiz-Primo, M. A.; Shavelson, R. J. (1996). *Journal of Research in Science Teaching, 33,* 569.
15. Ruiz-Primo, M. A., Shavelson, R. J., Li, M., Schultz, E. S. (2001). *Education and Assessment, 7*, 99.
16. Ruiz-Primo, M. A., P. (2004). *First International Conference on Concept Mapping.* Pamplona, Spain.
17. McClure, J. R.; Sonak, B.; Suen, H. K. (1999). *Journal of Research in Science Teaching, 36,* 475.
18. Ausubel, D. P.; Novak, J. D.; and Hariesian H. (1968). *Educational Psychology: A Cognitive View.* New York: Reinhart and Winston.
19. Piaget, J. (1926). *The language and thought of the child.* New York: Harcourt Brace.
20. Vygotsky, L.S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.
21. Herron, J. D. (1975). *Journal of Chemical Education, 52,* 146.

22. Herron, J. D. (1978). *Journal of Chemical Education, 55*, 165.

23. Bodner, G. (1986). *Journal of Chemical Education, 63*, 873.

24. Bodner, G; Klobuchar, M.; and Geelan, D. (2001). *Journal of Chemical Education 78*, 1107.

25. Taber, K.S. (2001). *Chemistry Education: Research and Practice in Europe, 2* (1), 43.

26. Novak, J. D. and Gowin, D. B. (1984). *Learning How to Learn.* Cambridge, MA: University Press.

27. Novak, J. D. (1990). *Journal of Research in Science Teaching, 27*, 937.

28. Naveh-Benjamin, M., McKeachie, W.J., Lin, Y.G., and Neely, R.K. (1998). *Contemporary Educational Psychology, 23*, 209.

29. Chang, K. E.; Sung, Y. T.; Chen, S. F., (2001). *Journal of Computer Assisted Learning 17*, 21.

30. Tsai, C. C.; Lin, S. S. J.; Yuan, S. M. (2001). *Journal of Computer Assisted Learning 17*, 72.

31. Rocha, F. E. L., Costa Jr., J. V., Favero, E. L., P. (2004). *First International Conference on Concept Mapping.* Pamplona, Spain.

32. Harrison, S. H., Wallce, J. L., Ebert-May, D., Luckie, D. B., P. (2004). *First International Conference on Concept Mapping.* Pamplona, Spain.

33. Luckie, D. B., Batzli, J. M., Harrison, S. H., Ebert-May, D. (2003). *International Journal of Learning, 10*, 1051.

34. Luckie, D.B., Harrison, S. H., Ebert-May, D., P. (2004). *First International Conference on Concept Mapping.* Pamplona, Spain.

35. Williams, K. S., Ebert-May, D., Luckie, D. B., Hodder, J., Koptur, S. (2004). *Front. Ecol. Environ., 2*, 444.

36. Edmonson, (1997). *Handbook of Academic Learning: Construction of knowledge.* San Diego, CA: Academic Press.

37. Francisco, J. S., Nakleh, M. B., Nurrenbern, S. C., Miller, M.L. (2002). *Journal of Chemical Education, 79*, 248.

38. Kinchin, I. M.; Hay, D. B.; Adams, A. (2000). *Educational Research 42*, 43.

39. Arneson, B. T., Lagowski, J. J., Abstracts of Papers, 223rd ACS National Meeting, Orlando, FL, CHED-174.

40. Stuart, H. A. (1985). *European Journal of Science Education, 7*, 73.

41. Ausubel, D. P. (2000). *The Acquisition and Retention of Knowledge: A Cognitive View.* Boston: Kluwer Academic Publishers.

42. Kinchin, I. M. (2000). *School Science Review 81*, 41.

43. Watson, A.H. and McCabe, T.J. (1996). *Structured Testing: A Testing Methodology Using the Cyclomatic Complexity Metric*. Gaithersburg, MD: National Institute of Standards and Technology.

44. Mintez, J. J., Wandersee, J. H., Novak, J. D. (1997). *Handbook of Academic Learning: Construction of Knowledge.* San Diego, CA: Academic Press.

Vita

Brian Todd Arneson was born in Brookfield, Wisconsin on August 17, 1974, the son of Mary Beth Arneson and Robert Michael Arneson. After completing her work at Menasha High School, Menasha, Wisconsin, in 1992, he enrolled at Michigan Technological University in Houghton, Michigan. He graduated from Michigan Technological University in May 1997 with a Bachelor of Science in chemistry. During the following year, he was employed as an analytical chemist for the Illinois State Geological Survey in Champaign, Illinois. In August 1998 he entered the Graduate School of The University of Texas. In August of 2003 he received a Master of Arts degree from The University of Texas in the field of chemistry.

Permanent address: 8300 Renton Dr., Austin, Texas 78757.

This dissertation was typed by Brian T. Arneson.